Molecular studies towards understanding the role of Prep1 (Pbx Regulation Protein 1) in segmental expression of Hox proteins in the vertebrate hindbrain

Thesis

UNRESTRICTED.

Elisabetta Ferretti

Molecular studies towards understanding the role of Prep1 (Pbx Regulation Protein 1) in segmental expression of Hox proteins in the vertebrate hindbrain.

Thesis submitted in partial fulfillment of the requirements of the Open University for the degree of Doctor of Philosophy in Molecular and Cellular Biology

2003

Sponsoring establishment:

Open University, London

Director of study

Prof. Blasi Francesco

External supervisor

Prof. Krumlauf Robb

Collaborating establishment:

DIBIT

Department of Biological and Technological Research

San Raffaele Scientific Institute

Milan, Italy

Submission date: 31 August 2003

Award date: 26 February 2004

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

TABLE OF CONTENTS

Acknowledgment

First, I would like to thank all the people that supported me during the past four years of PhD.
I thank my boss Prof Francesco Blasi for giving me the possibilty to work in his laboratory, supporting my project and helping me to become an independent researcher.
I am extremly grateful to my external supervisor, Prof. Robb Krumlauf for suggestions, support and for the extraordinary collaborative work without which this thesis would been much less. I would like to thank Miguel and Nobue that I met at MRC in London.
A special thought to the people of the Stower Institute: Paco, Paul, Xinghao, Debs, Valerie, Reme, Stefan, Leta, Takaoski, Lienn and Heather who accompanied me during my four months in Kansas City. They were wonderful in welcoming me so warmly with friendship and support.
I also thank EMBO who made this visit possible.
Another acknowledgment goes to Dr Francesco Argenton and to his laboratory for the wonderful collaborative work performed on the zebrafish project.
Special thanks to Vincenzo Zappavigna for the fruitful discussions and to the member of Blasi's laboratory for their patience to manage my impulsive character.
I have no words to express my gratitude to Roberta, Orla and Daniela for helpful discussions especially during the writing of this thesis.
A special thought is dedicated to the Scu-ba group (Cristiano, Barbara, Alexia and Fabrizio) and to my housemate Cristhian, they were precious in showing me the richness and beauty of things outside the laboratory (living in the sea).
Finally, thanks to my family and specially to my brother Roberto. All the work, experiments and efforts producted in this thesis are dedicate to him, because he does not believe that a thesis could take so much effort and time!
GRAZIE

Betta

Declaration

This PhD thesis is submitted in partial fulfilment of the requirement of the Open University.

This thesis has been composed by myself and has not been used for any previous degree from this or any other University.

The experiments presented in this thesis are original data and were carried out at the Molecular Genetcis Department, DIBIT H.S. Raffaele Scientific Institute, Milan under the guidance of Prof. Francesco Blasi.

Some of the presented data were obtained in collaboration. At the end of each chapter I have indicated the contribution made by other people to the work presented

ABBREVATIONS

A/P	antero-posterior
Aa	amino acids
AML	acute myeloid leukemia
ANR	anterior neural ridge
Antp	Antennapedia
ARE	autoregulatory element
b1- ARE	Hoxb1 autoregulatory element
BA	branchial arch
bHLH	basic Helic-Loop-helix
BMP	bone morphogenetic proteins
bp	base pair
cDNA	complementary DNA (DNA print of mRNA)
CNS	central nervous system
d.p.c	days post coitum
D/V	dorso-ventral
Dpp	decapentaplegic
EC	embryonic carcinoma cells
EMSA	electrophoretic mobility shift assay
ES	embryonic stem cells
Exd	extradenticle
Fgf	fibroblast growth factor
FISH	fluorescent in situ hybridization
GC	granulose cells
GFP	green fluorescent protein
h.p.f	hours post fertilization
HAT	histone acetyltransferase
HD	homeodomain
HDAC	histone deacetylase
HH stage	Hambuger-Hamilton stage
Hh	hedgehog
HM	Meis homeodomain
HOM-C	Homeotic selector genes in Drosophila
Hox	Homeotic selector genes in mammals
HPIP	hematopoietic Pbx interacting protein
HR	Homology Region
Hth	Homothorax
Kb	kilobase pair
KDa	kilo Daltons
Lab	labial
LC	Lower complex
MAP	microtubuline associated protein
Meinox	Meis and Knox proteins

Meis	Murine ectopic insertion site
MLT	murine leukemia virus
MRNA	messenger RNA
NC	neural crest
Neo	neomycin resistance gene
NES	Nuclear Export Signal
NLS	Nuclear Localization Signal
NMR	nuclear magnetic resonance
ORF	open reading frame
PAGE	poly-acrylamide gel electrophoresis
PBC	Pbx, Exd and Ceh20
Pbx	Pre-B cells homeobox gene
PCR	Polymerase
PFA	paraformaldehide
PGK	phosphate-buffered saline
PH	Pbx/Hox binding site
PKA	protein kinase A
PM	Pbx/Meinox binding site
Prep1	Pbx regulation protein
PRS	Pbx responsive site
PVDF	polyvinyldifluoride
R1	repeat 1 of b1-ARE
R2	repeat 2 of b1-ARE
R3	repeat 3 of b1-ARE
r4	rhombomere 4
RA	retinoic acid
RARE	retinoic acid responsive element
RT	room temperature
RT-PCR	reverse transcriptase-PCR
SDS	sodium dodecyl sulphate
TALE	three amino acid loop extension
UC	Upper complex
UEF3	uPA enhancer factor 3
UPA	urokinase plasminogen activator
UTR	untranslated region
Wg	wingless

ABSTRACT

The homeotic genes of the *Hox* cluster specify segmental diversity along the body axis of vertebrates. This segmental modulation of target genes produces morphologically distinct segments termed rhombomers which define the fate of cells on the antero-posterior body axis of embryos (McGinnins and Krumlauf, 1992). Members of the Hox gene family are sequence-specific DNA binding transcription factors which show high conservation within the DNA binding homeodomain. Consequently, the DNA binding properties of Hox proteins are very similar, raising the question of how proteins with equivalent biochemical properties can achieve *in vivo* distinct regulatory effects and define precise developmental programs (Ghering *et al.*, 1994; Lu *et al.*, 1996). The mechanisms by which the different Hox protein and their cofactors generate those diverse fates remain unclear (Mann, 1995). Both of evidence supports the model where the DNA binding sites that directly interact with Hox/PBC heterodimers determine which member of the Hox protein family will bind and thereby regulate a given target element (Rauskolb *et al.*, 1993; Popperl *et al.*, 1995; Di Rocco *et al.*, 1997). Furthermore, the Meinox protein family (Prep, Meis and Hht) form specific heterodimers with the PBC proteins acting as cofactors. Thus, the identification of Meinox proteins that can regulate the PBC proteins introduces an additional level of complexity in the regulation of *Hox* protein function.

We have isolated two members of the Meinox family: Prep1 and Prep2. We also showed that Prep1 and Prep2 represent a new sub class of Meionox subfamily, distinct from Meis. We analyzed the expression of Prep1 both in adult and in embryonic murine tissues showing that Prep1 is ubiquitously expressed in developing embryos as well as in adult mice.

We also showed that Prep1 is expressed early in the mouse embryo and is already present in the cytoplasm of murine oocytes, suggesting it involvement in the initial stages of embryonic development.

Prep1 is always found in a DNA-binding complex with members of the Pbx protein family, whose identity varies among different tissues and during embryogenesis. Thus, the Prep1-Pbx heterodimers show a tissue and development-dependent composition, suggesting specific regulation properties for the different complexes.

We found that Prep1, Pbx1 and Pbx2 are present in the hindbrain at the time when Pbx is controlling *Hox* expression. This colocalization is compatible with a role for Prep1 as a regulator of Pbx-Hox interaction and function *in vivo*. Furthermore, we have demonstrated that r4-specific *Hoxb1* and *Hoxb2* and r5/r6 specific *Hoxa3* enhancers are complex elements containing separated Pbx/Hox (PH) and Pbx/Meinox (PM) binding sites. We have also shown that the PM-PH motifs are elements present in several *Hox* enhancers and conserved in different species from fly to mammals. Moreover we have shown that both the PM and PH sites are required for the *in vitro* formation of a Prep1-Pbx-Hox ternary complex. We also showed that, *in vivo*, the PM site of *Hoxb2* enhancer is essential for Hoxb2 expression in r4. On the contrary, mutations in the PM site of *Hoxb1* and *Hoxa3* enhancers do not alter Hoxb1 r-4restricted and Hoxa3 r5/r5 restricted expression (Ferretti *et al.*, 2000; Manzanares *et al.*, 2001).

We propose that the additional PH sites in *Hoxb1* and *Hoxa3* enhancers modulate the activity of those enhancers and may explain the different in vivo behaviors. We demonstrated, that *in vitro,* one of the additional PH site (R2) in the *Hoxb1* enhancer exerts an inhibitory effect on the ternary complex formation. We have further demonstrated that the PM-PH elements from *Hoxb1* and *Hoxb2* enhancers are sufficient to direct the r4-specific expression of the reporter gene in r4 of electroporated chicken embryos.
Moreover, *in vivo* like *in vitro,* the presence of R2 exerts an inhibitory effect by blocking the r4-restricted expression of the reporter gene in the chicken embryos

Finally we have inactivated Prep1 both in mouse and in zebrafish showing that Prep1 has an essential role in embryonic development.

Prep1 -/- mice show an embryonic lethal phenotype, dieing in uterus at 16.5 d.p.c. Gross morphologic analysis of Prep1-/- embryos revealed massive subcutaneous edema, generalized pallor diminished vascularization, smaller livers and abnormally orientated forelimbs. All these abnormalities observed in Prep1-/- embryos and those presented by Pbx1-/-embryos are similar, suggesting a genetic interaction between these two factors.

Inactivation of *prep1.1* in zebrafish results in a lethal phenotype. *prep1.1* morphants show severe alteration in cranial-facial chondrogenesis resulting from of the incapacity of neural crest to differentiate into chondroblast.

CHAPTER 1

INTRODUCTION

Different cell types make different sets of proteins even though they have identical genomes. How is it possible to obtain different phenotypes with the same genotype? In vertebrates, cell fate (specification and determination) is not rigidly preordained, instead distinct identities are acquired through a cells specific combination of transcription factors, surface receptors, diffusible molecules with the ability to select specific pathways and form precise target connections.

The integration of cell extrinsic and cell intrinsic signals is mediated primarly through the fine regulation of gene expression._The control of gene expression takes place at several levels: differential gene transcription, selective RNA processing, selective RNA translation, differential protein modification and protein degradation.

There are proteins that regulate transcription (transcription factors) by recognizing specific DNA sequences through discrete DNA-binding domains. These domains are in general relatively small, less than 100 amino acid residues. A common DNA-binding domain in prokaryotes is the helix-turn-helix structure that recognizes and binds specific regulatory DNA regions. In eukaryotes there are diverse DNA-binding domains including the zinc finger, leucine zipper and helix-turn-helix structures.

The combination of specific DNA sequences and transcription factors generates a unique control module. Different combinations of sets of control modules lead to the specific regulation of each gene. This strategy allows the production of a high number of combinations starting from a low number of modules. According to the enhanceosome theorem, enhancer regions of eukaryotic genes mediate the assembly of stereospecific multiprotein containing complexes, i.e. transcription factor complexes (Tjian and Maniatis, 1994). Therefore efficient gene expression depends upon a series of interactions between a set of DNA-binding proteins and their corresponding DNA sequences (regulatory regions: promoters, enhancer and silencer). Each cell has the same DNA control sequences, but not every cell has the same set of DNA-binding proteins. Additional regulatory events control the accessibility of transcription factors to the DNA at the level of chromatin structure, nuclear localization, transcription, translation, post-translational, modification, *etc*. Most cellular functions, such as differentiation, embryonic development and response to the environment, depend on the presence/absence and/or activation/inactivation of transcription factors.

Transcription factors can be divided into two classes: general factors, which are ubiquitous and are required for basal transcriptional activity and specific transcription factors that are present in a specific cell type at a specific time and are responsible for the expression of cell type-specific and time-specific

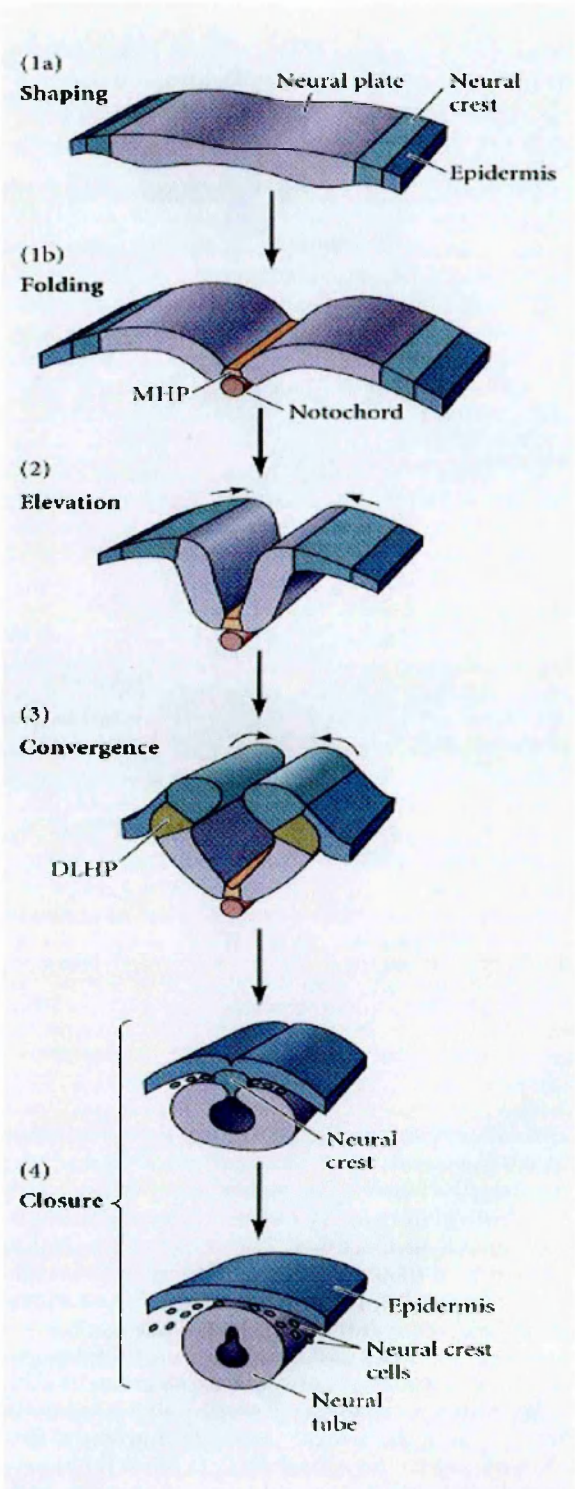


Figure 1.1
Schematic representation of neural formation the chick embryo. The neural tube from flat sheet of neural tissue, termed the neural plate (1a). The neural plate bends ventrally at the medial neural hinge point (MHP) and is anchored to notocord (1b). The neural folds rise up and join at the dorsal midline, the dorso lateral hinge point (DLHP), producing a closed neural tube with the epidermis (4) (Figure taken from Gilbert et al., 1996).

proteins. An important example of a biological event that requires a stringent and precise control of gene expression is the establishment of body axes and organogenesis during embryonic development. In particular, specific transcription factors called homeotic proteins play a crucial role in controlling embryogenesis and specifying antero-posterior axis formation. In vertebrates, homeotic proteins are actively expressed in mesoderm, ectoderm and endodermal tissues that are undergoing extensive cell proliferation, migration and cell-cell communication associated with early development and organogenesis.
Homeotic genes show overlapping domains of expression in the developing embryo. The pattern of expression of these homeotic genes is called the Hox code. More than one homeotic gene can be expressed in any given cell, which may cooperate or compete for common binding sites within the regulatory sequences of their target genes.

Development of the main regions of the Central Nervous System (CNS)

The CNS has traditionally been subdivided into two main regions: the brain and spinal cord. The brain and the spinal cord are organized to allow the input of primary sensory information, the analysis and processing of information and the transmission of appropriate responses to the information (Butler and Hodos, 1996). The CNS originates from the neural plate (Fig 1.1), a cytologically homogeneous layer of neuroectodermal cells that form the dorsal surface of the gastrula stage embryo. During the generation of the nervous system, the first step is the specification of a homogeneous layer of ectodermal cells into the neural plate and the epidermis (Fig 1.1). The neural plate is first distinguishable in the mouse embryo after 7.5 days post coitum (d.p.c.). Studies in the *Xenopus Leavis* embryo suggest that an essential contribution to neural induction comes from a region located in the blastopore lip, termed the Organizer. The Organizer tissue produces morphogenetic molecules, like Bone Morphogenetic Protein (BMP), which repress the neural and promote the epidermal cell fate (Tanabe and Jessel, 1996). Other potential organizer molecules expressed by axial mesoderm are Follistatin, Noggin and Chordin. These three proteins exhibit no obvious common structural features but each can interact and antagonize the actions of members of the BMP protein family thus inducing neural tissue formation (Wilson *et al.*, 1995).

Thus, the induction of neural plate involves the inhinition mechanisms of TGF-β-like signals in order to promote epidermal and repress neural differentiation.

For example, biochemical studies in *Xenopus Leavis* have shown that Noggin binds directly to BMP2 and BMP4, thus preventing them from binding to their receptors and resulting in neural induction (Hansen *et al.*, 1997).

The axis induction organizer is called the Hensen's node in chick, the shield in zebrafish and the node in mouse (review Lemaire and Kodjabachian, 1996; Camus and Tam, 1999). Other molecules that have an essential role in the

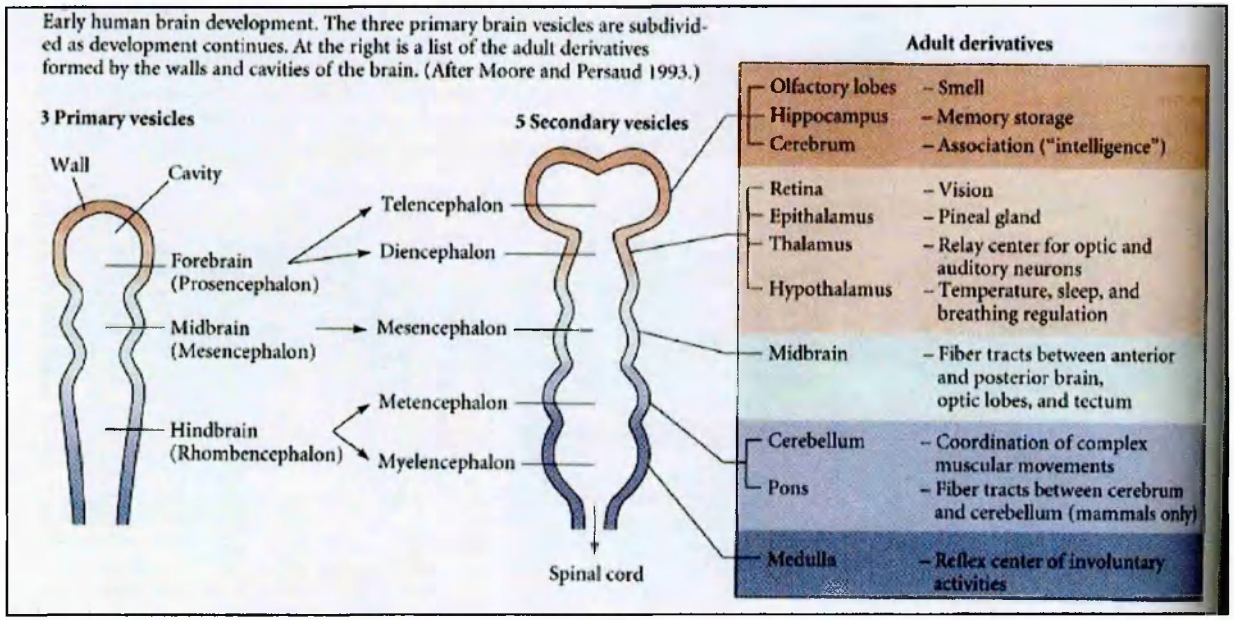


Figure 1.2
Schematization of early human brain development, starting from the formation of the four brain vesicles (forebrain, midbrain, hindbrain and spinal cord) to the derived adult brain tissues. At the rostral end of the neural tube two vesicles, the telencephalon and diencephalon, make up what is considered the forebrain. The telencephalon divides in two symmetrical anterior vesicles that form the cerebral cortex. Posteriorly, the mesencephalon develops into midbrain. The rhomboencephalon, or hindbrain, give rise most anteriorly to the metencephalon and posteriorly to the myelencephalon. The spinal cord makes up the posterior region of the CNS. Each brain vesicles gives rise to different adult structures (Figure from Moore and Persaud, 1993)

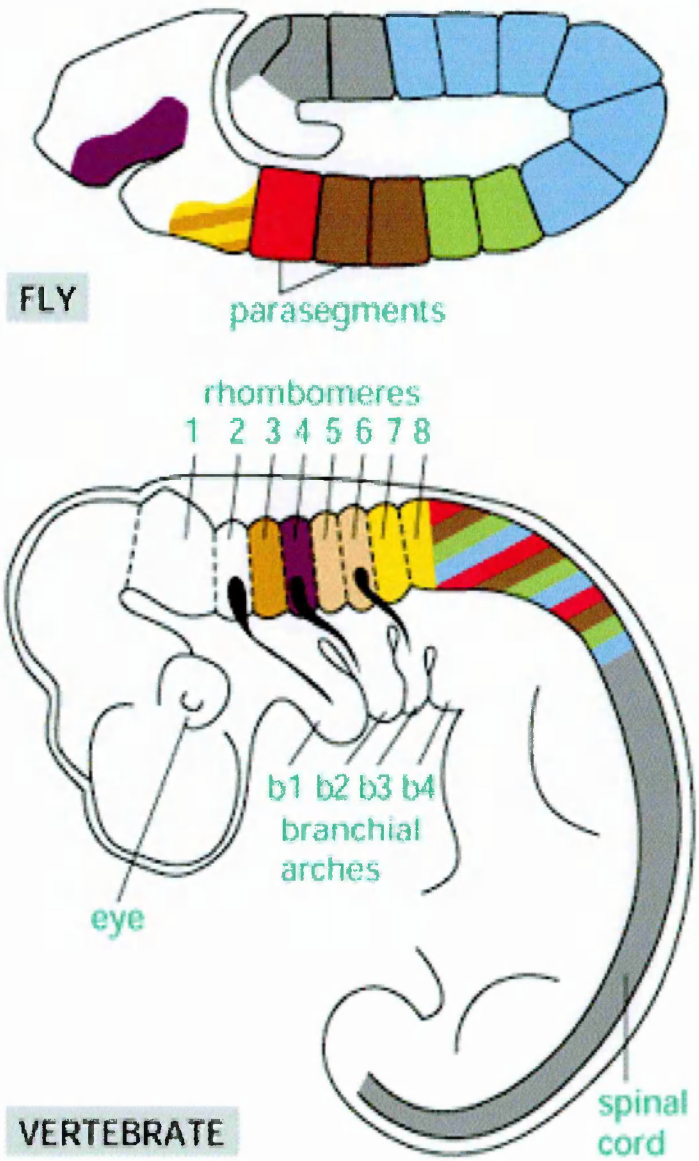


Figure 1.3
Schematization of fly (Drosophila) and vertebrate (mouse) development. The hindbrain is the unique vertebrate structures showing a a evident sub-division in segments similar to the Drosophila embryo. Parasegment in fly and rhombomers in mice represent transient structures which a specific and unique identity

organizer function are the signaling molecules transforming growth factor-β (TGF-β) and Wnt1 (Conlon *et al.*, 1994).

Around 8-8.5 d.p.c, the neural plate dramatically changes shape and its lateral edges elevate to form the neural fold (Fig 1.1). Subsequently, the neural folds fuse along the dorsal midline and form the neural tube. This step generates the CNS and its components, neurons and glial cells. Cells in the region of this fusion form a specialized group of cells known as the neural crest cell (NC) (see later). Delamination of neural crest cells generates different cells types, including Schwann cells and sensory neurons.

Neural inducing factors and modifiers produced during gastrulation by the mesoderm establish an initial antero-posterior (A/P) axis in the developing neural tube. The induction of anterior neural plate differentiation involves the inhibition of two other signaling molecules fibroblast growth factor-β (FGF-β) and Wnt1 (Lamb and Harland, 1995). The exact nature of the signaling molecules involved is not known.

At 8.5-9 d.p.c. the neural tube in mouse embryo forms four vesicles: forebrain (prosencephalon), midbrain (mesencephalon), hindbrain (rhombencephalon) and spinal cord (Fig 1.2). At the rostral end of the neural tube two vesicles, the telencephalon and diencephalon, originate from the forebrain, while secondary bulges (the optic vesicles) extend laterally from each side of the developing forebrain and later form the eyes. Posteriorly, the metencephalon, or midbrain, will generate an important region (isthmic region) that connects the forebrain and hindbrain.

Developing hindbrain generates two vesicles: the metencephalon and myelencephalon. Seven transient compartments termed rhombomers characterize the hindbrain. Finally, the caudal region of the neural tube corresponds to the spinal cord, extending from the brain to the tip of the tail.
Thus, from the early stages of the development the CNS appears morphologically divided into regions that will later develop into distinct adult structures. Each vesicle appears to use a specific set of developmental mechanisms to establish the A/P axis, whereas the mechanisms of dorso/ventral (D/V) patterning are similar.

Hindbrain

The most caudal region of the developing brain is the hindbrain, which will generate later the medulla and the pons in the adult (Fig 1.2). The medulla and pons form the fourth ventricle of the brain and, together with the midbrain, are often referred to as the brainstem. The brainstem contains a conglomerate of cell groups that form a complex network termed the reticular formation. This network is involved in higher order behavior such as respiration, circulation and wakefulness. The reticulospinal neurons of the brainstem are important in connecting the brain to the spinal cord to control locomotion.
Closely related with the pons both geographically and functionally is the cerebellum, a cortical structure responsible for coordinating movement, posture and balance. Patterning of the hindbrain involves segmentation of the neural tube along the anterior posterior (A/P) axis. During development (at 9 d.p.c. in mouse), the hindbrain is transiently divided into a series of seven lineage-restricted compartments termed: the rhombomeres (r1 to r7) and the lateral

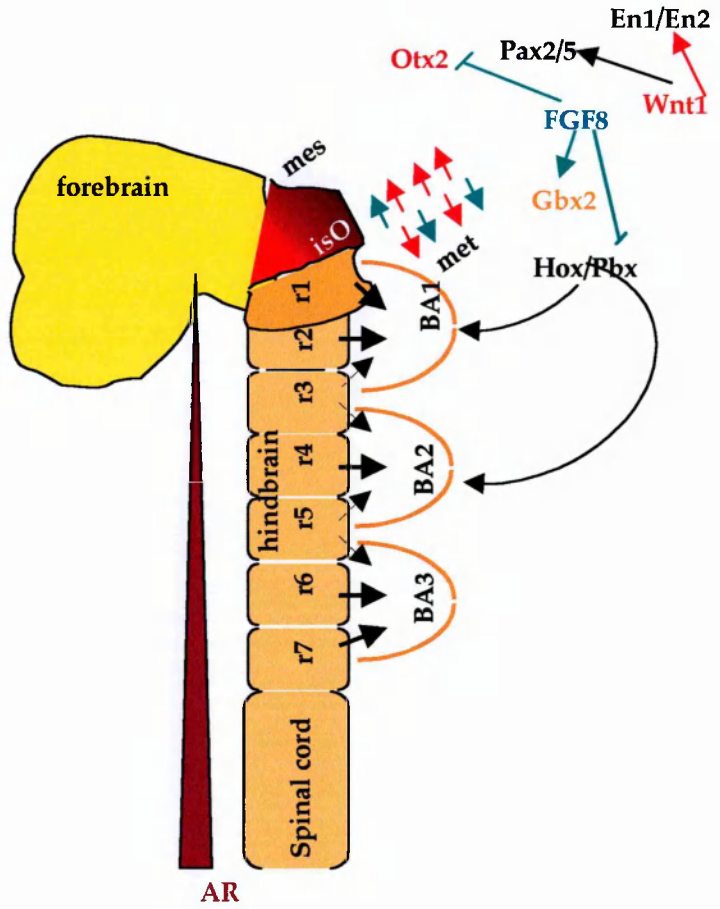


Figure 1.4

Schematization representation of mid-brain-hindbrain patterning in mouse embryo (9.5 d.p.c). The positioning and maintenance of the establishment of the isthmic organizer (isO) involved different genes, including Otx1, Otx2, En1, En2 Pax2 , Pax5 and Gbx2 factors
FGF8 (blue) positively controls Gbx2 expression in the metencephalon (Met) and repress Otx2 in the mesencephalon (Mes).
FGF8 (blue) and Wnt1 (red) factors expressed in the isthmus are required for positioning and maintenance of the border of Otx2/Gbx2. Thus may there are responsible for the organizing activity of the isthmus.
Moreover, FGF8 represses the anterior border of Hox genes. Hox genes and their cofactors Pbx are involved in the establishment of rhombomere identity.

mesoderm forms a series of branchial arches (BA), which are populated by neural crest cells (Fig 1.3; Lumsden and Keynes, 1998). Rhombomeres can function as developmental compartment structures and limited cell movement and mixing occurs between neighboring rhombomeres. As a consequence, each rhombomere appears to be formed by cells that have the same lineage and follow the same developmental fate based on the local information within each compartment (Guthrie and Lumsden, 1991). Each rhombomere adopts a distinct set of cellular and molecular properties that are different from those of its neighbors. In these compartments the cells acquire distinct properties, thus permitting a degree of autonomy during the period of cell specification (Garcia-Bellido *et al.*, 1973; Lawrence, 1989). Rhombomeres acquire early unique morphological and molecular information, that confers a specific rhombomere identity. The concept of positional identity has important implications for development. The result is that a cell, or group of cells, in the embryo acquires states related to a specific position at a specific time and this determines later development. The morphological boundaries between the rhombomeres correspond to the boundaries of expression of the Hox proteins, which are evolutionarily conserved regulators of segment identity (see later) (Wilkinson *et al.*, 1989). The expression pattern of the murine *Hox* genes suggests that certain combinations of *Hox* genes specify a particular region of the anterior-posterior axis (the *Hox* code) (Hunt and Krumlauf, 1991; Kessel and Gruss, 1991). Therefore, different and specific combinations of Hox genes, and their partners, characterize different regions (rhombomeres) of the hindbrain. Furthermore, auto-cross-and para-regulative mechanisms with Hox proteins and their cofactors (PBC and Meinox proteins) are required to maintain specific Hox pattern expression (Fig 1.3 and see later for details).

Knockout or transgenic mis-expression studies of Hox genes in mice have shown that individual rhombomeres may be lost or transformed to different identities. This confers a functionally essential role to the *Hox* genes in the patterning of the hindbrain, suggesting that the segmentation of the hindbrain appears as a transient state, presumably required initially to establish a framework of A/P pattern within the hindbrain.

In addition, the hindbrain segmentation is essential to define pathways of cranial neural crest migration, which in turn gives rise both to peripheral nerves and to skeletal elements (see later).

Midbrain

The isthmic region (mid/hindbrain junction) is located in the midbrain, beyond the anterior limit of *Hox* gene expression (Fig 1.4). Transplantation studies between chick/quail have shown that the isthmic region has the ability to change the fate of the host cells, inducing the formation of midbrain structures and maintaining the same A/P polarity as the transplanted tissues (Fig 1.4) (Wassef and Joyner, 1996). In addition, the host cells in different A/P positions in the neural tube respond differentially to the same isthmic signals, suggesting the presence of a molecular gradient in the midbrain. Interestingly the *engrailed* genes (*En*) are expressed in a decreasing manner and correlate to the A/P. Two homeodomain containing proteins, Otx2 and Gbx2, have a specific gene expression border within the isthmus (Fig 1.4). The early event that defines the

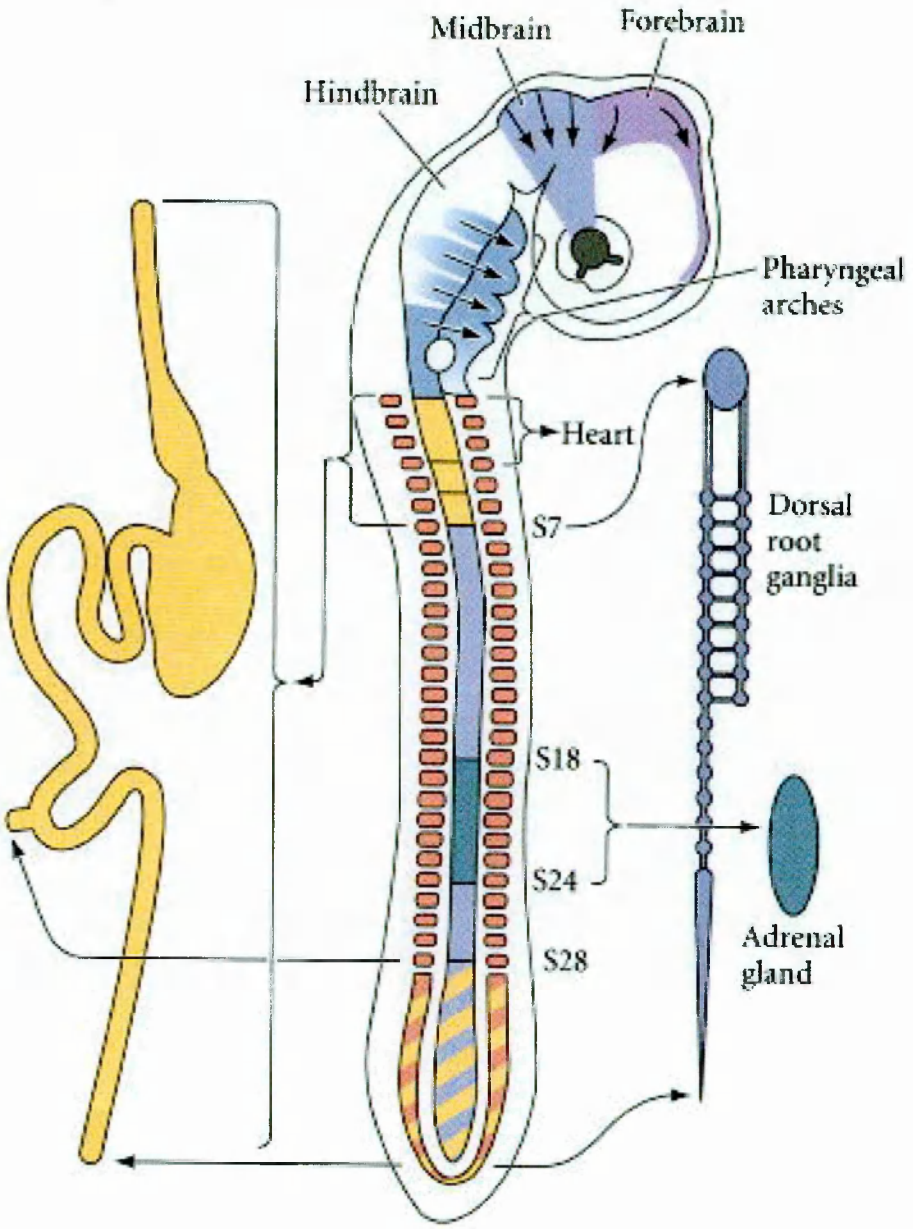


Figure 1. 5

Schematization of migration of neural crest cells.The cranial neural crest migrate into the branchial arches and face to form the bones and cartilage of face and neck. They also produce pigment and constitute the sensory part of same cranial nerves. The vagal neural crest and the sacral neural crest form the parasympathetic nerves of gut.The cardiac neural crest are involve in making the division between the aorta and the pulmonary artery. Neural crest cells of the trunk make the sympathetic neurons and are also involve in formation o medulla portion of the adrenal gland (from Le Douarin 1982).

Mesencephalon/Metencephalon (Mes/Met) A/P axis is the division of the neural plate into anterior positive Otx2 and posterior Gbx2 domains. Tissues that are Gbx2 positive will form cerebellar structures, whereas Otx2 positive cells will produce midbrain structures. The signals that control Otx2 and Gbx2 expression domains remain to be identified, but retinoids and FGF8 seem to be involved in this process (Fig 1.4). In the next stage other signaling molecules such as Wnt1, which is essential for the elaboration of midbrain and rostrally for hindbrain development, are involved. Wnt1 is secreted in the rostral isthmus, where it is implicated in the expression control of En1, En2, Pax2, Pax5, Otx1 and Otx2 transcription factors (Fig 1.4). FGF8 is secreted in the caudal isthmus where it positively controls Gbx2 metencephalic expression and represses Otx2 and the rostral Hox genes (Fig 1.4). All these proteins contribute to the patterning of midbrain development. Double Pax2/Pax5 knockout mice show deletion of the isthmus (Urbanek *et al.*, 1994). Mice lacking En1, En2 or Wnt1 have a similar phenotype, showing deletion of midbrain and cerebellum (Wassef and Joyner, 1996).

Forebrain

The forebrain, the most rostral part of the brain, consists of two major regions: the diencephalon and the telencephalon. The telencephalon forms the cerebral hemispheres and the striatum.

In contrast to hindbrain and midbrain patterning, where restricted patterns of gene expression are linked to segmentation or to the activity of signaling region, the mechanisms of mouse forebrain patterning are still unknown. It has been proposed that the forebrain, like the hindbrain, is composed of metameric units called prosomers. However, the absence of clear cell lineage restriction boundaries and evident repeated expression pattern does not support this theory. Signaling proteins, as in the isthmus-midbrain, could generate patterning in the forebrain. The anterior boundary of the neural plate anterior, neural, ridge (ANR) expresses FGF8 and this suggests that this factor could be a possible candidate for such a function (Fig 1.4). Moreover, local patterning from neural and non-neural sources contributes to increase the regional complexity of the forebrain. One of the most important diffusible molecules is Sonic hedgehog (Shh), which is involved not only in the D/V but also in the A/P patterning of the telencephalon.

Neural crest cells

The neural crest cells originate during the closure of the neural tube in the region that connects the neural tube and the epidermis, but they then migrate extensively to generate different cell types, including neurons and glial cells of the sensory sympathetic and para sympathetic of the neuron system, melanocytes and much of the skeletal and connective tissue components of the head (Fig 1.5). While all the neuronal crest cells can form melanocytes, neurons and glia, only the cranial neural crest cells (see below) are able to produce

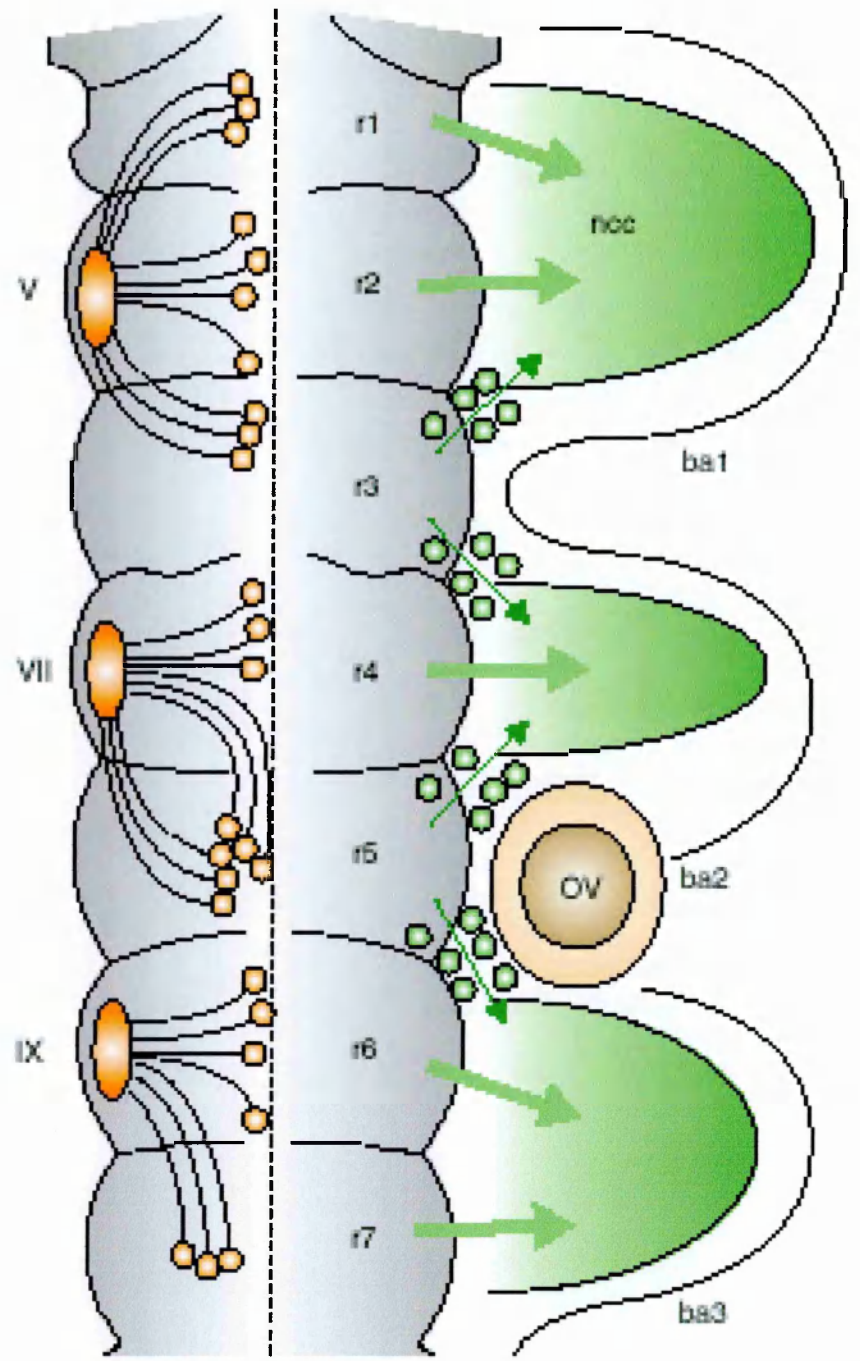


Figure 1.6
Segmental organization of the hindbrain, motor nerves and pathways of cranial crest cells. The branchiomotor nerves exit from the hindbrain only from the even numbered segments (orange ovals) to innervate their peripheral targets. Neural crest cells migrate laterally from r1, r2, r4, r6 and r7 (green arrows) into the branchial arches (ba). R3 and r5 generate neural crest cells that migrate to branchial arches joining the even numbered rhombomeres. ncc: neural crest cells; OV: otic vesicle; V: trigeminal motor nerve; VII:facial motor nerve; IX: glossopharyngeal motor nerve. (taken from Trainor and Krumlauf, 2001).

Pharingeal arch	skeletal elements (neural crest plus mesoderm)	Arches arteries (mesoderm)	Muscles (mesoderm)	Cranial nerves (neural tube)
1	Incus and malleus (from neural crest); mandibole, maxila, and temporal bone regions (from crest dermal mesenchyme),	Maxilar branch of the carotid artery (to the ear, nose and jaw).	Jaw muscles; floor of mouth; muscles of the ear and soft palate	Maxillary and mandibular division of trigeminal nerves (V).
2	Stapes bone of the middle ear; styloid process of temporal bone; part of hyoid bone of neck (all from neural crest cartilage).	Arteries to the ear region:cortico-tympanic artery (adult); stapedial artery (embryo).	Muscles of facial expression; jaw and upper neck muscles	Facial nerve (VII)
3	Lower rim and greater horns of hyoid bone (from neural crest).	Common carotid artery; root of internal carotid	Stylipharyngea us (to elevate the pharynx)	Glossopharynge al nerve(IX).
4	Laryngeal cartilages (from lateral plate mesoderm).	Arch of aorta right subclavian artery; original spout of pulmonary arteries	Constriction of fharynx and vocal cords	Superior laryngeal branch of vagus nerve (X).
6	Laryngeal cartilages (from lateral plate mesoderm).	Ductus arteriosus; roots of definitive pulmonary arteries	Intrinsic muscles of larynx	Recurrent laryngeal branch of vagus nerve (X).

TABLE I from Larsen 1992

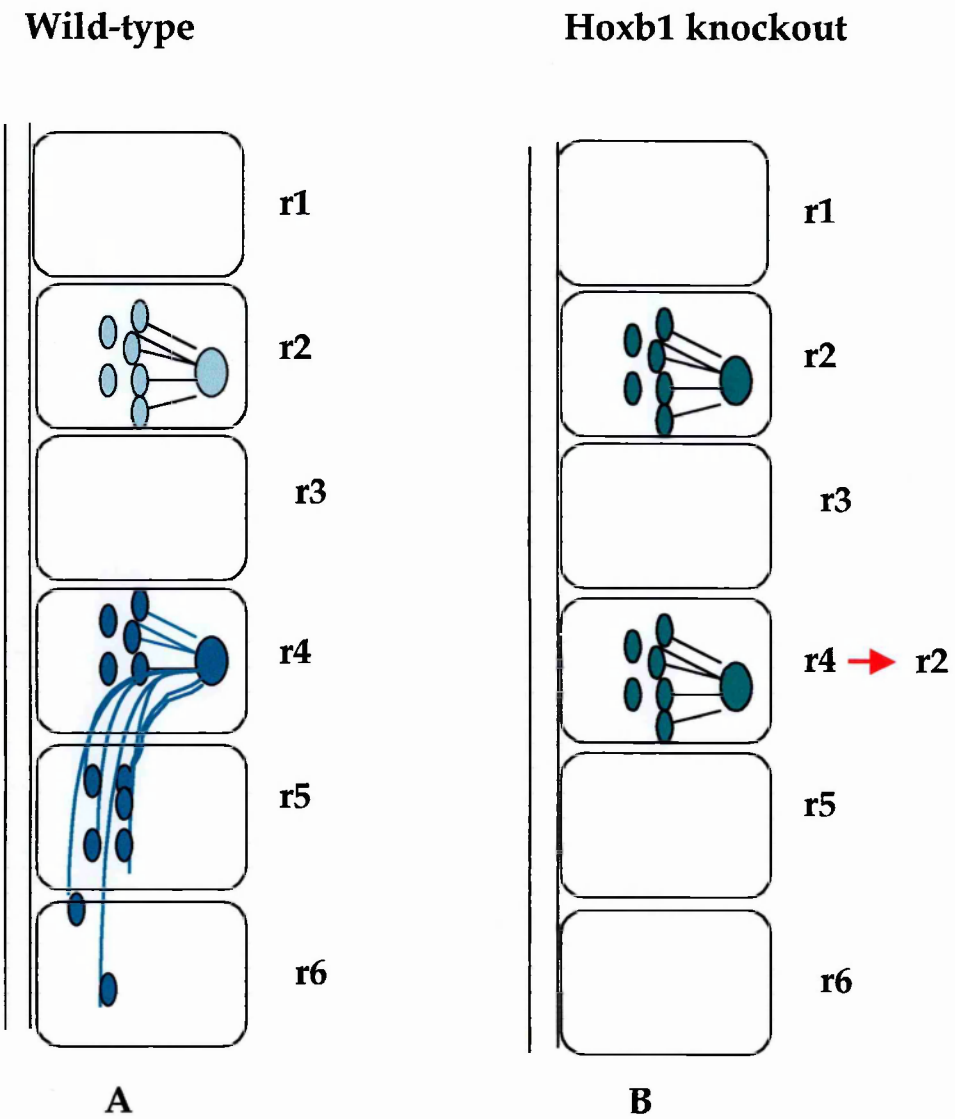


Figure 1.7
Rhombomere 4 change proper identity in r2. In Hoxb1 mutant embryos, motor neurons generated in r4 fail to migrate caudally and direct laterally in similar manner o trigeminal motor neuron. In Hoxb1-/- mice the cells in r4 generate trigeminal rather than facial neuron (modified from Studer *et al.*,m 1995)

cartilage and bone. The fate of neural crest cells depends principally on where they migrate and settle (see tableI).

The branchial arches (BA) are metameric structures that develop in a cranial-caudal sequence, and are formed by epithelial tissues covering mesenchymal bars of paraxial mesoderm and cranial neural crest cells.
There are six pairs of BA (BA5 is only rudimentary) and in most mammals it is difficult to discriminate BA4-6. The neural crest cells can be subdivided into four main functional domains.

Cranial neural crest

Cranial neural crest cells play a critical role in the construction of the vertebrate head (Fig 1.5 and Fig 1.6). They migrate to the branchial arches and contribute to specific cranial sensory ganglia, form bone and cartilage derivatives and influence muscle patterning and branchiomotor targets (Le Douarin and Ziller, 1983; Noden, 1986).

During early embryo development the hindbrain exerts a crucial influence on craniofacial morphogenesis, in part through its ability to generate cranial neural crest. Neural crest cells emerge from the dorsal margin of each rhombomere. The hindbrain segmental organization is critical for the proper spatial organization of the cranial ganglia, branchiomotor nerves and pathways of cranial neural crest migration. Hindbrain derived neural crest cells migrate laterally from r1, r2, r4, r6 and r7 into the branchial arches (Fig 1.6). Cranial neural crest cells from rhombomere 1 and 2 migrate into the first branchial arch (mandibular) producing the jawbone and the incus and malleus bones of the middle ear. In addition these cells generate the bones of face (Lumsden *et al.*, 1991). Cranial neural crest cells from rhombomere 4 migrate to the second branchial arch (hyoid) forming the hyoid cartilage of the neck (Lumsden *et al.*, 1991). Cranial neural crest cells from rhombomeres 6 and 7 migrate into the third branchial arch (visceral) to form the thymus, parathyroid and thyroid glands (Chisaka and Capecchi, 1991). Cranial neural crest cells originating from rhombomeres 3 and 5 migrate rostrally and caudally to join the stream arising from the even numbered rhombomers (Fig 1.6) (Lumsden *et al*, 1991). Surgical manipulation studies indicate that even-numbered rhombomers appear to exert a repressive effect upon the production of the neural crest by odd-numbered rhombomeres, probably by inducing cell death by apoptosis (Graham *et al.*, 1993).

Cranial neural crest cells originating in the forebrain and midbrain contribute to the frontonasal process, palate and mesenchyme of the first branchial arch. Cranial neural crest cells migrating into the fourth and sixth branchial arch produce the neck bones and muscles (the fifth BA degenerates in humans). Neural crest cells that migrate from r1 and r2 contribute to the formation of trigeminal ganglion, while those from r4 the facial and vestibulo-acustic ganglia and those from r6 the superior ganglion of the IXth nerve (Fig 1.6). Gain and loss of function analysis demonstrates a functional role of Hox genes in neural crest patterning and head development. Altering *Hox* expression pattern, in addition to changing rhombomere identity, modifies neural crest cell specification (Krumlauf, 1994; Hunt *et al.*, 1991). For example, in *Hoxa2* knock mice second (hyoid) arch NC derivates have been replaced by a mirror –image duplication of a subset of first arch-like structures, suggesting the the second BA

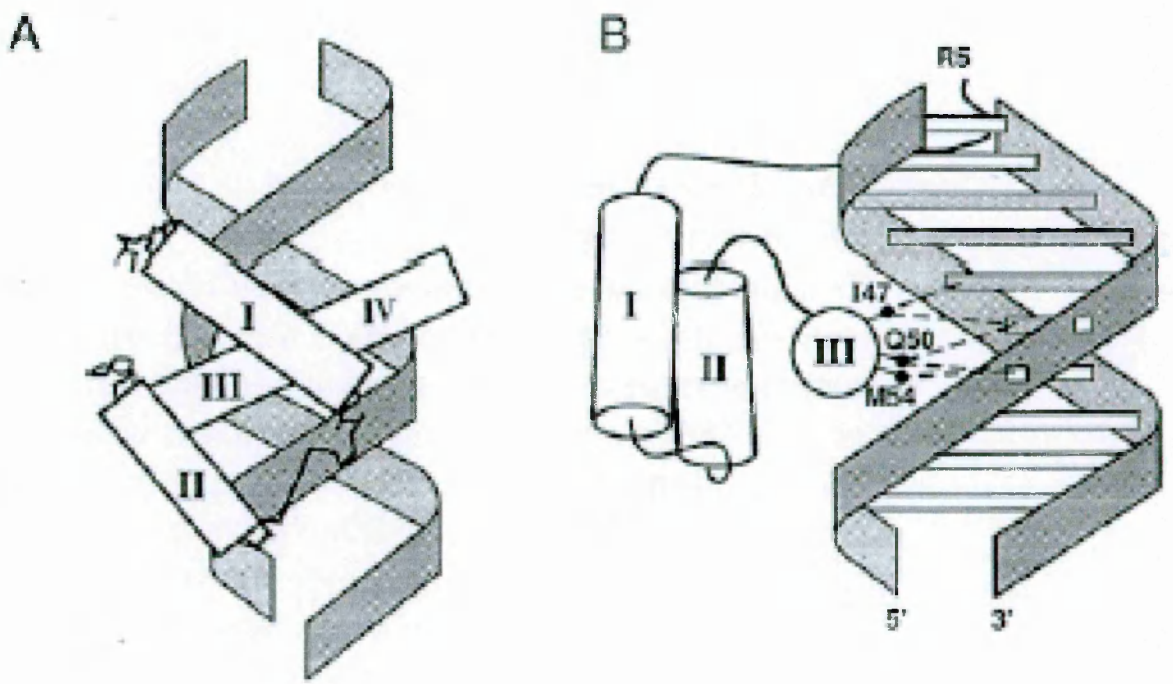


Figure 1.8
Schematic representation of three helices (I, II, III shown as cylinder) forming the homeodomain. A) Frontal view (B) side view. Both pictures show that the third helix of the homeodomain is inserted into the DNA major groove, while the N-terminus arm R5 contacts the minor groove. The helix IV is a direct extension of helix III.

derivates are homeotically transformed in those normally derived derived from r1+r2 crest programme (Kontges and Lumsden, 1996).

The ability of neural crest cell to maintain the *Hox*-code from their region of origin is suggested by grafting experiments (Prince and Lumsden, 1994; Couly *et al.*, 1996). However, cell transposition studies have shown that the cranial neural crest cells have a high degree of plasticity in *Hox* expression pattern (Salvidar *et al.*, 1996; Trainor and Krumlauf, 2000). Thus cranial neural crest cells can be reprogrammed in specific branchial arch environments, and cranial mesoderm seems to be the source of a signal to maintain or change the *Hox* pattern (Trainor and Krumlauf, 2000).

Trunk neural crest

The cells migrating dorso-laterally become the pigment-synthesizing melanocytes, while the cells that migrate ventrolaterally arrive in the sclerotomes. The cells that remain in the sclerotome form the dorsal root of the ganglia sensory neurons, while the cells that pass ventrally through the sclerotome differentiate into the sympathetic ganglia, the adrenal medulla and the nerve cluster surrounding the aorta (Fig 1.5).

Vagal and sacral neural crest

These cells generate the parasympathetic ganglia of the gut (Fig 1.5).

Cardiac neural crest

These cells are located between the cranial and trunk neural crest and can develop into different tissues such as melanocytes, neurons, cartilage and connective tissue. In addition they form the muscle connective tissue wall of the large arteries and are critical in making the division between the aorta and the pulmonary artery (Fig 1.5).

Homeodomain proteins

The homeodomain proteins play instructive roles in regulating embryonic development, therefore in generating the body plan in higher organisms (McGinnis and Krumlauf, 1992; Krumlauf, 1994). The term "homeotic" derives from the Greek word "homeo" meaning "similar" and was originally used for fruit fly (Drosophila) mutants in which another replaced one structure. The first homeotic transformation identified in Drosophila was the Antennapedia mutation, in which the antennae were replaced by legs. The homeotic genes specify the identity of a particular body segment (Mc Ginnies *et al.*, 1984; Scott *et al.*, 1984).

The homeotic genes contain a conserved 180 bp sequence called the homeobox, encoding a 60 amino acid motif termed the homeodomain. This homeodomain acts as a DNA-binding domain, recognizing specific DNA sequences (Mc Ginnies *et al.*, 1984; Scott *et al.*, 1984). The homeodomain has been found in a large variety of proteins and it is evolutionarilly conserved throughout the two kingdoms. In Drosophila, only the genes that belong to the

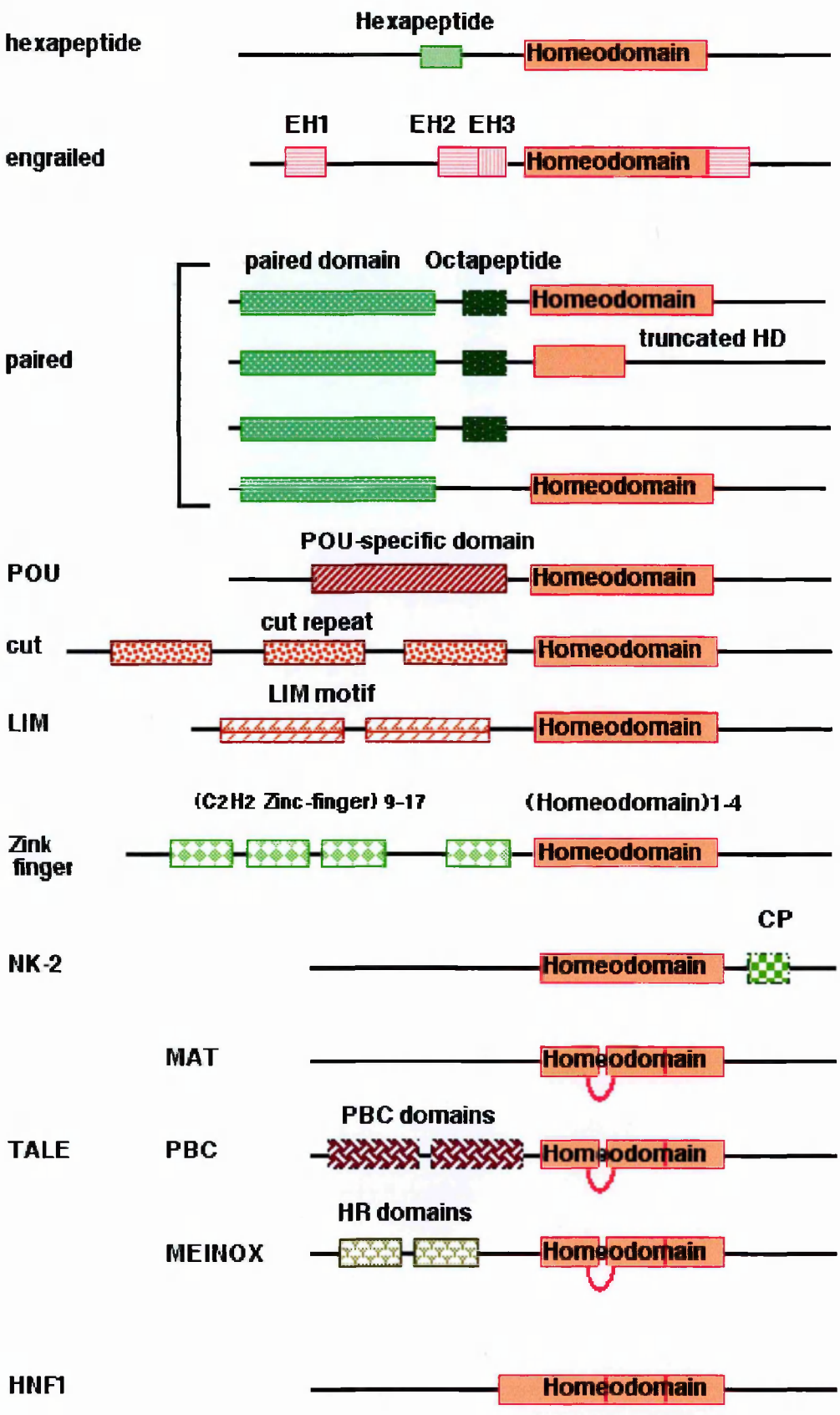


Figure 1.9
Schematic representation of Dispersed Super class homeodomain containing proteins (modified from Gehring *et al.*, 1994).

HOM-C cluster and give rise to homeotic transformation are called homeotic. All other genes are termed homeodomain containing proteins.

In vertebrate homeotic, only the genes that belong to the *Hox* cluster are called, not all those gene produced homeotic transformation when mutated. However, cofactors of homeotic proteins, belonging to the Three Amino acids Loop Extension (TALE) family (see later) also show homeotic transformation when they are inactivated (Popperl *et al.*, 2000; Selleri *et al.*, 2001; Waskiewicz *et al.*, 2002).

Structure of the homeodomain

As mentioned, the homeodomain containing proteins have a genetic regulatory function. They act as transcription factors regulating target genes in a precise spatial and temporal pattern.
2D NMR analysis of the Antennapedia homeodomain in complex with its corresponding DNA provided the initial structural information on these domains. Subsequently, these studies were completed by the characterization of the crystal structure of homeodomain-DNA complexes of engrailed, the yeast MAT α2 and even skipped (Wolberger *et al.*, 1991a/b).

Structural data show that the homeodomain folds into three α-helices (Fig 1.8). Helices I and II are connected by a hexapeptide and arranged in an antiparallel manner, whereas helix III and its flexible extension (also termed helix IV) are perpendicular to the first two. Helices II and III generate a helix-turn-helix conformation, similar to the prokaryotic DNA-binding motifs. Helix III is the recognition helix, and specifically contacts the major DNA groove. All homeodomains contain four invariant residues, Trp 48, Phe49, Asn 51 and Arg 53, in the recognition helix. The two polar amino acids provide strong interactions with DNA. Residue Asn 51 forms a hydrogen bond with one of the adenine residues of the ATTA consensus-binding site. The other two residues are part of a hydrophobic core and are required for the recognition helix to reach the correct position with respect to the entire homeodomain. The hydrophobic core is highly conserved, suggesting that different homeodomains can fold into a similar three-dimensional structure. The N-terminal region of the homeodomain is a flexible amino terminal arm capable of binding the minor groove of DNA with Arg5 that contacts a thymine in the consensus-binding site. Position 9 of helix III has a crucial role in differentiating between similar recognition sites. Thus, it is sufficient to replace Lys 9 in the Bicoid homeodomain with a Glu residue to change the binding specificity towards *Antennapedia* target sites. The N-terminal is the most divergent region of the homeodomain, but is not sufficient to confer specificity to the homeodomain proteins. In fact, homeodomain proteins with different functions can have identical amino acids in position 9, suggesting that others mechanisms have developed to increase the binding specificity of homeodomain containing proteins (Qian *et al.*, 1992).

Classification of homeodomains

Homeodomain sequences can be subdivided into at least 20 different classes on the basis of several criteria; sequence similarity in the flanking regions, organization into gene clusters, homology with other sequence motifs and position of introns (Fig 1.9). On the basis of sequence similarity and

chromosomal clustering (which reflect evolutionary relationships) the homeodomain containing proteins can be classified into two super classes.
1. Genes of the Complex Super class are clustered homeotic genes including the Drosophila *HOM-C* and Mammalian *Hox* genes. This super-class has been subdivided into six classes on the basis of the corresponding Drosophila genes: *labial (lab), proboscidea (pb), deformed (Ddf), sex combs reducer (Scr), Antennapedia (Antp), Abdominal-B (Abd-B).* Drosophila homeotic genes and their mammalian counterparts (*Hox* genes) share high homology in the homeodomain and, with the exception of the *Abd-B* class, contain a hexapeptide motif with the consensus sequence (IYPWMK), located N-terminal to the homeodomain. All the genes are arranged in clusters and are subjected to the colinearity effect (see below).
2. Genes of the Dispersed Super class are not clustered and encode proteins with a divergent homeobox that also have additional protein domains (Fig. 1.9). Several sub-classes have been identified and they are listed below.

The hexapeptide class: Cad, Emx and TCL/NEC, contain a hexapeptide motif that is also present in the clustered homeotic proteins (Gehring *et al.*, 1994).

The engrailed (en) class is highly conserved in evolution and is characterized by four highly conserved EH protein segments outside the homeodomain (Figure 1.9, Gehring *et al.*, 1994).

The paired class contains two DNA binding domains, a homeodomain and a region of 128 amino acids termed the paired domain. Some proteins contain both DNA binding domains, some only the paired domain and some a truncated homeodomain. Moreover, some proteins include an additional octapeptide motif (Fig. 1.9, Gehring *et al.*, 1994).

The Pou class includes several ubiquitous transcription factors such as Pit1, the octamer motif-binding transcription factors Oct1 and Oct2. In addition to the POU-homeodomain, the POU proteins contain an additional DNA-binding domain termed the POU-specific domain (Gehring *et al.*, 1994). This bipartite DNA binding domain is divided by a flexible linker that permits to the protein to adopt various monomer configurations on DNA (Fig. 1.9, Scholer et al., 1991; Verrijzer and Van der Vliet, 1993).

The cut class is a small group of proteins containing three copies of an 80 amino acid "cut" repeat upstream of the homeodomain (Fig. 1.9, Gehring *et al.*, 1994).

The LIM class is characterized by the presence, in addition to the homeodomain, of LIM motifs, which are quite variable. The LIM motif is a protein-protein interaction motif, and is required for the interaction with other homeodomain containing proteins like POU and Otx proteins (Fig. 1.9, Gehring *et al.*, 1994).

The zinc finger class contains both a homeodomain and a well-characterized protein-protein interaction domain termed the zinc finger. The zinc fingers are of the C_2-H_2 type and most closely related to the "gap" genes Kruppel and hunchback of *Drosophila* (Fig. 1.9).

The NK-2 class contains a conserved peptide (CP) downstream of the homeodomain (Fig. 1.9).

The TALE (Three Amino acids Loop Extension) class is characterized by an atypical homeodomain that has three extra amino acids in the homeodomain that loop out between helices I and II. Usually the three aminoacids are proline-tyrosine-proline (Burling, 1997). This turn is often followed by serine or

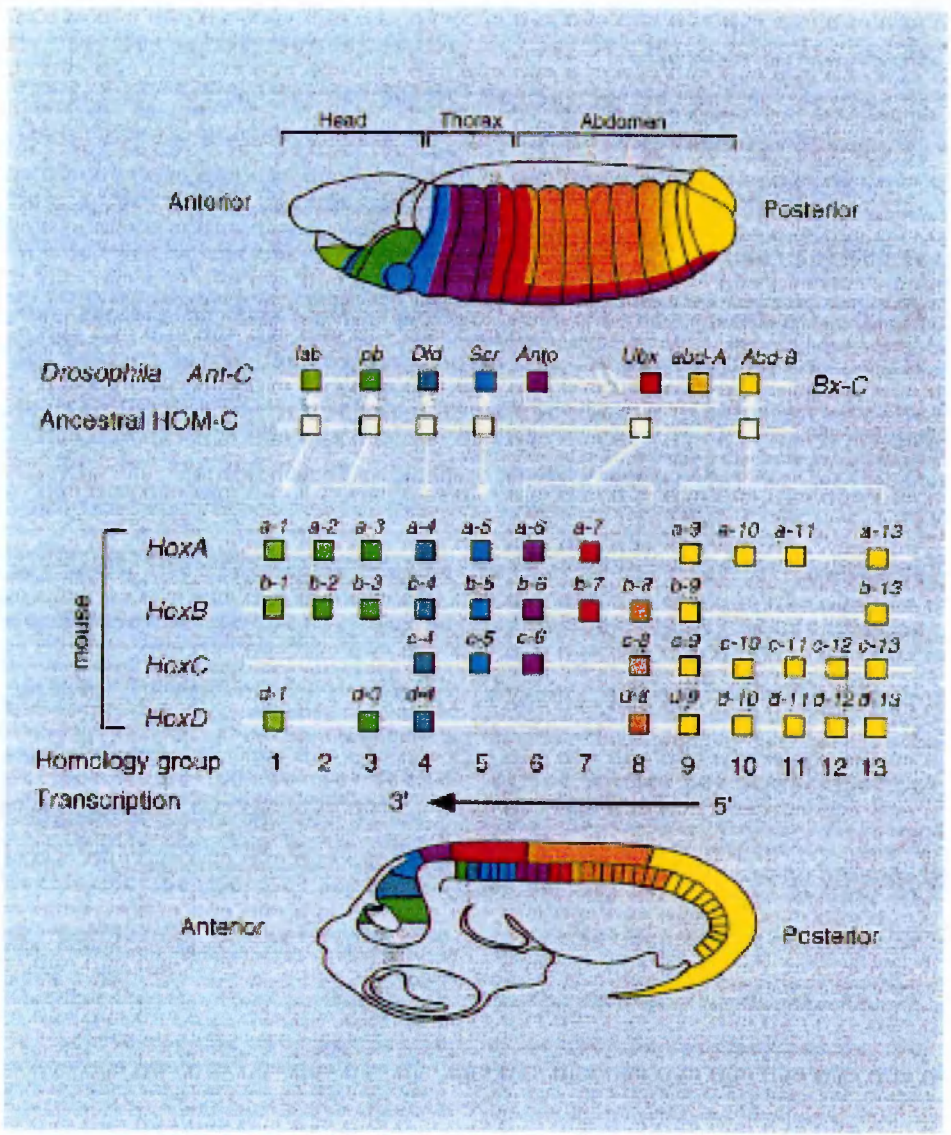


Figure 1.10
Evolutionary conservation of homeotic gene organization and transcriptional expression in Drosophila and mice . Similarity between the HOM-C cluster chromosome 3 and embryonic development expression. The scheme also shows the colinearity of expression throughout the posterior-anterior axis of the Hox genes in both Drosophila larvae and the murine embryo.

threonine and several acidic residues (Burlin *et al.*, 1997). The function of this loop is still unknown (Fig. 1.9).
The yeast mating type MAT α1 and α2 were the first atypical homeodomain genes isolated. Subsequently, atypical homeodomain genes were isolated from Drosophila, nematodes and mammals and they include the PBC and the Meinox proteins. The TALE proteins will be discussed in detail below.

The rat liver transcription factor HNF1 contains a more extreme atypical homeodomain of 81 amino acids.

Drosophila *HOM-C*

It has already been mentioned that the homeotic genes, both in Drosophila and in vertebrate animals, play a fundamental role in determining the segmental identity during embryonic development (Mc Ginnies *et al.*, 1984; Scott *et al.*, 1984).

During embryo development, the homeotic genes act after the maternal genes (*bicoid, hunchback, nanos*), the gap genes and the paired rule genes (Mc Ginnies *et al.*, 1984; Scott *et al.*, 1984). They define the A/P polarity and the segmental pattern of the embryo, larva and adult by generating 14 morphological units, termed segments, along the antero-posterior axis of the embryo. The *HOM-C* genes establish cell identity within the segment. Each segment is characterized by a unique combination of homeotic genes that confers identity to that segment (Fig 1.10).

Mutation in *HOM-C* genes does not change the number of segments, but rather their identity (Carpenter *et al.*, 1993; Dolle' *et al.*, 1993; Mark *et al.*, 1993).
There are two regions of Drosophila chromosome 3 containing most of the homeotic genes. One region, the *Antennapedia* complex, contains the homeotic genes *labial (lab), Antennapedia (Antp), sex combs reducer (scr), deformed (dfd)* and *proboscidea (pb)* (Fig 1.10). The labial and deformed proteins are required for the identity of the head segments, while the sex comb reducer and Antennapedia are important to specify thoracic segments. The proboscidea protein acts in the adult organism.

The second region is the *bithorax* complex that contains three genes: *ultrabithorax (ubx)*, required for thoracic segment identity, *abdominal A (abdA)* and *abdominal B (abdB)*, responsible for the identity of abdominal segments. Not only mutations in the coding regions, but also mutations in the *cis* regulatory regions generate homeotic transformations. For example, a mutation of the regulatory region of *Ultrabithorax* gene transforms the third thoracic segment into a second thoracic segment (*i.e.* halteres into wings).

The peculiar hallmark of the *HOM-C* complex is the correlation between the physical order of genes along the chromosome and their expression/function along the A/P axis of the embryo (Fig 1.10). This property is called colinearity.

The chromosome region containing *Antennapedia* and *Bithorax* complexes is termed the homeotic complex or *HOM-C*.

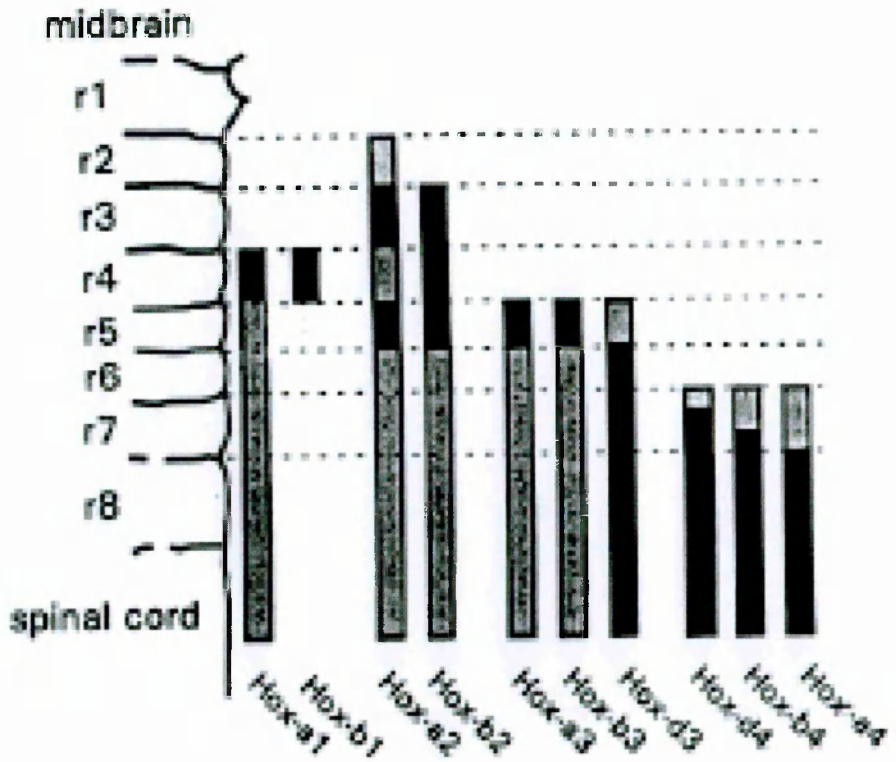


Figure 1.11
Schematic representation of the segmental boundaries of *Hox* genes expression along the murine hindbrain at 9.5 d.pc.(from Keynes and Krumlauf, 1994).

HOM-C binding specificity

The *HOM-C* genes control the choice between different developmental pathways. For example, the *Antp* gene governs the choice between pathways resulting in a leg or in an antenna. The Antp homeodomain is able to bind to DNA regulatory elements controlling a unique set of leg-promoting and antennal suppressing target genes (Botas *et.al.*, 1993).
Antp binds with high affinity to the sequence [C/T][C/A]ATTA, but this sequence is very abundant in the Drosophila genome. Therefore, nature must have evolved mechanisms modifying Antp DNA-binding selectivity that control its *in vivo* functions. Control could be exerted at the level of chromatin structure and accessibility to some sequences. In addition, the interaction of Antp with other factors could modulate its binding specificity. Several studies have implicated extradenticle (Exd) as a HOM-C cofactor. Exd has a fundamental role in defining segmental identity in the Drosophila embryo. Exd mutations induce homeotic transformation without altering the expression pattern of the *HOM-C* genes, leading to the conclusion that Exd in some way acts functionally in parallel with homeotic genes (Pfeiffer and Wieschaus, 1990). Exd is subjected to a regulation of nuclear-cytoplasm localization, which affects its ability to interact with nuclear HOM-C proteins (this topic will be discussed in further detail below) (Abu-Shaar and Mann, 1996; Aspland and White, 1997).

Hox Genes

Expression and function of *Hox* genes

The *Hox* genes are evolutionary homologues of the homeotic factors that were originally identified in Drosophila (*HOM-C* genes) and have been found in all vertebrates. High similarities in genomic structure and organization suggested that *Hox/HOM-C* clusters arose by duplication and divergence from a common ancestral cluster (Fig 1.10) (McGinnis and Krumlauf, 1992; Krumlauf, 1994). Human and mice genomes contain 38 *Hox* genes, organized into four different chromosomal complexes (termed A, B, C and D). Because the different complexes are the result of duplications, *Hox* genes with the same rank in different complexes are evolutionary relatives, and are therefore called paralogues (Fig 1.10). The peculiar feature of *Hox* gene expression is that the rostro-caudal and temporal order (from the 3'to the 5') of expression follows the order of the genes in the chromosome, a property termed colinearity. Thus, the physical order of *Hox* genes on the chromosome correlates with the order of anterior boundaries of expression of each gene along the anterior-posterior axis of the embryo (Fig 1.10) (Lewis, 1978; Duboule and Dolle', 1989; Graham, 1989).
Hox gene nomenclature emphasizes this feature: within each complex the genes are numbered in order of rostral to caudal expression domains, maintaining rank alignments according to sequence similarity (Fig 1.10). The rostro-caudal order has few exceptions (for example, *Hoxb2* is expressed rostral to *Hoxb1*), and is surprisingly well conserved in different species (Fig 1.11). *Hox* gene expression begins early in embryogenesis, typically during the formation of the primitive

streak in mice (7.5 d.p.c). *Hox* gene expression is very dynamic and can change rapidly during development. For example, *Hoxa1* is highly expressed in r3 and r4 at 8 d.p.c. and then recedes caudally. At 8.5 d.p.c. *Hoxa1* expression is no longer detectable in the hindbrain.

Hox gene expression involves the nervous system (hindbrain), somite derived structures (prevertebral column) limbs, gut, and gonads tissue (Dolle' *et al* 1989; and 1991; Izpisua -Belmonte *et al.*, 1991; Yokouchi *et al.*, 1991; Roberts *et al.*, 1995).

Hox genes are expressed along the neural tube, neural crest, paraxial mesoderm and surface ectoderm starting from the anterior boundary of the hindbrain through to the tail. The pattern of *Hox* gene expression in the ectoderm and branchial arches at specific positions along the A/P axis is similar in the neural tube and cranial crest. The correlation in terms of *Hox* gene expression between hindbrain and emigrating neural crest cells suggests that the *Hox*-code may extend to the branchial arches. However, *Hox* gene expression pattern in the branchial arches is more complex and the combination of *Hox* genes expressed by neural crest cells can be altered by the environment into which they migrate (Saldivar *et al.*, 1996).

Thus, *Hox* genes are known to play important roles in specifying regional differentiation of several tissues in vertebrates including neuroectodermal tissues, limbs, gut, and mesoderm (Dolle' *et al.*, 1989; Izpisua-Belmonte *et al.*, 1991; Kessel and Gruss, 1991). For example, within the hindbrain, modification of the pattern of Hox expression alters the regional pattern of neuronal differentiation. Ectopic expression of Hoxb1, for example, affects the facial motoneurons (Studer *at al.*, 1998).

Loss and gain of function analyses have shown that *Hox* genes control the specification of rhombomere identities (homeotic transformation). In mouse *Hox* knockouts, individual rhombomeres are frequently lost or partially transformed to more anterior identities. *Hox* gene deletion usually affects the most anterior regions in which the gene is expressed, leaving posterior structures relatively unaffected. This phenomenon is known as posterior dominance or posterior prevalence (Duboule, 1995).

Hoxa1 knockout mice fail to form rhombomeres r3-6 (Rossel and Capecchi *et al.*, 1999; Gavalas *et al.*, 1998). *Hoxb1* -/- mice show alteration of migration of motoneurons (Fig 1.7). In *Hoxa2*-null mutants, the rhombomere 1/2 boundary is absent (Davenne *et al.*, 1999). Moreover, the size of rhombomeres 2 and 3 is reduced and rhombomere 1 is enlarged, resulting in posterior expansion of the cerebellar territory (which is normally derived mainly from rhombomere 1) and the neurons that are normally derived from r2 are missing (Gavalas *et al.*, 1997). However phenotypes observed in a *Hox* single mutant are often subtle, due to the combinatorial *Hox* code that characterizes most rhombomeres, and to the partial redundancy among *Hox* paralogues (Greer *et al.*, 2000). In fact, a severe phenotype is associated with loss of both *Hox-1* genes (*Hoxa1* and *Hoxb1*), which is visible as a facial paralysis (Studer *at al.*, 1998; Rossel and Capecchi, 1999).

By contrast, strong effects are observed when *Hox* genes are expressed abnormally in anterior regions where they normally are not expressed. This can result in transformation of anterior structures in others that are normally more posterior. For example, the over-expression of *Hoxa2* in the first chick branchial arch leads to transformation of the first arch into a second arch cartilage, thus

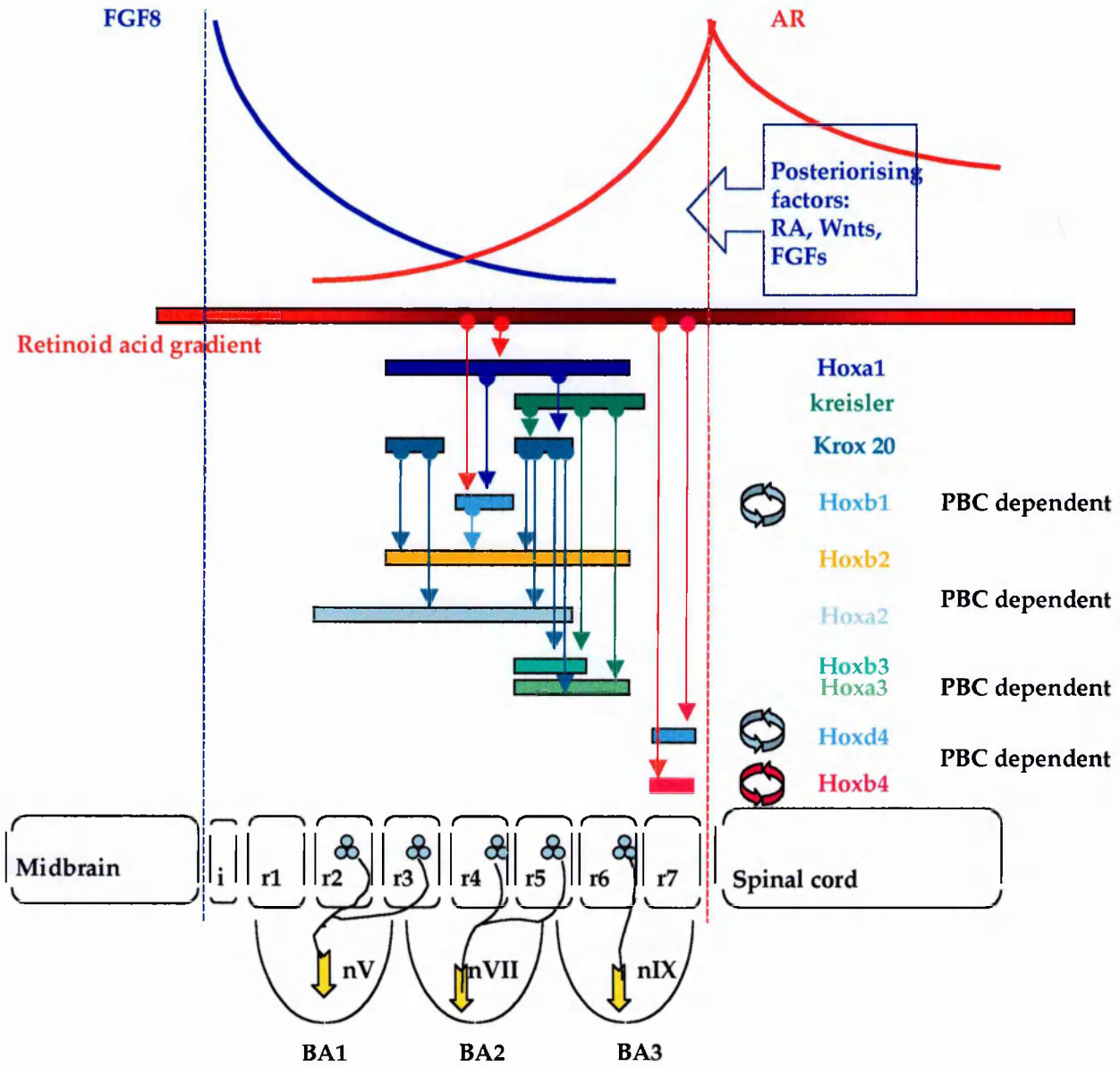


Figure 1.12
Schematic representation of regulator molecules acting on the *Hox* genes activation cascade along the antero-posterior axis of the developing embryo. Retinoid acid and FGF8 form opposite gradiens that control *Hox* expression. Regional *Hox* expression establishes the positional identity and information for hindbrain patterning and neural crest cell migration. The big arrow indicates the factors showing posteriorization effects, while the curves represent the gradient of morphogen molecules present along the hindbrain. The neural crest cells (yellow arrows) emerging from hind brain migrate to the first three branchial arches (BA) and contribute to form the sensitive part of nerves (nV, nVII and IX). The circles correspond to an autoregulatory mechanisms.

transforming Meckel cartilage into quadrato cartilage (Grammatopoulos *et al.*, 2000). When *Hoxb1*, which is expressed at highest levels in r4, was mis-expressed only in r2, the r2 motor axon projection was reorganized to resemble an r4 motor axon (Bell *et al.*, 1999).

Hox regulators

Understanding how the expression of *Hox* genes is established and maintained is of critical importance, since studies in many species have shown that shifts in the expression boundaries can lead to transformations and alterations of segmental identity (McGinnis and Krumlauf, 1992; Trainor *et al.*, 2000; Moens and Prince 2002).

Using reporter genes it has been possible to reconstruct patterns of expression for many of the 3′ members of the *Hox* complexes (anterior *Hox*). However, the analysis of the regulation of the more 5′ *Hox* (posterior *Hox*) has been quite complex.

The early activation of the 3′*Hox* genes involves different upstream regulators such as Krox20, Kreisler and retinoid receptors. The maintenance of these protein's expression domains is mediated through auto-and cross-regulatory interactions of the *Hox* genes themselves. On the contrary, the 5′ *Hox* gene cluster does not contain regulatory elements known to bind retinoic receptors, Kreisler or Krox20, indicating that other mechanisms must be involved. Mechanisms proposed for the regulation of the 5′ *Hox* genes include sharing enhancers (Nonchev *et al.*, 1996; Gould *et al.*, 1997, Sharpe *et al.*, 1998; Manzanares *et al.*, 2001) and response to a graded balance between two morphogens like FGF8 and retinoic acid (RA) (Mathis *et al.*, 2001; Streit *et al.*, 2000; Bel-Vialar *et al.*, 2002). Recently in chicken embryos it has been demonstrated that the 3′ *Hox* genes respond to RA but not to FGF8 treatment while the 5′ *Hox* genes are responsive to FGF8 upon activation of Cdx (Bel-Vialar *et al.*, 2002). Thus extracellular signaling molecules (RA and FGF) play a fundamental role in modulating *Hox* gene expression.

The signaling molecules (RA and FGFs).

A variety of experiments have shown that retinoic acid can induce several developmental abnormalities and interfere with the normal establishment of the *Hox* gene expression pattern. Moreover, RA induces changes in cell fate and the associated *Hox* genes expression in a manner consistent with the principle of colinearity (*Hox* genes at 3′ respond more rapidly and at a lower RA concentration than 5′ *Hox* gene) (Simeone *et al.*, 1990). For example, treatment of early mouse embryos with RA results in homeotic transformation of vertebrae (Kessel and Gruss, 1991; Conlon, 1995).

The highest level of the RA morphogen occurs at the spinal cord/hindbrain boundary, and gradually diffuses in both directions. The highest concentration of FGF8 occurs at the isthmus and spreads posteriorly into the hindbrain. Therefore, the hindbrain patterning is modulated by two opposing gradients (Fig 1.12).

This gradient theory appears to be in contrast with the observation that the more anterior *Hox* genes are more sensitive to RA (Fig 1.12). However, this

inverse gradient can be explained when both synthesis and degradation of retinoids is taken into consideration. The RA effective distribution pattern in the embryo correlates with the opposing action of the two metabolic enzymes RALDH-2 (which converts retinaldehyde into RA), and CYP26 (which oxidatively inactivates it), suggesting that the embryo is actively removing RA from the anterior neuroepithelium (Swindell *et al.*, 1999; Berggren *et al.*, 1999). It has yet to be shown that a posterior-to-anterior gradient of RA exists during hindbrain patterning, but the expression of RALDH-2 and CYP26 is suggestive of a source and sink that would be required to set up such gradient.

Similar studies in frog and chicken embryos suggest that FGF and RA work independently to define distinct regions along the A/P axis, and that the Cdx homeobox gene (vertebrate homolog of caudal Drosophila gene) is responsible for transducing the FGF signals (Isaacs *et al.*, 1998; Bel-Vialar *et al* 2002).

In the hindbrain, *Hoxa1* and *Hoxb1* enhancers contain RA response elements (RAREs) required for their initial activation. Multiple RAREs contribute to the expression of *Hoxb1*. One is specific for the neuroectoderm, whereas a second 3′ RARE is responsible for expression in the endoderm and developing gut. The analysis of double retinoic, acid, receptor (RAR) mutant mice has suggested that retinoids are not only required to establish the early anterior limit of *Hoxb1*, but they also participate in restricting hindbrain domains (Studer *et al.*, 1994; Marshall *et al.*, 1994).

Upstream Hox regulators

Once RA activates the *Hox* cascade in the hindbrain, how are the rhombomere-restricted *Hox* domains established and maintained? Several transcription factors play important roles in this process.

Krox 20 is a zinc finger containing protein that is specifically expressed in r3 and r5 and it regulates the expression of *Hoxa2* and *Hoxb2* in r3 and r5 *via* conserved binding sites in their enhancers. Targeted disruption of *Krox 20* results in the lack of r3 and r5 and in the formation of partially fused r2-r4-r6 territories.

Similarly, Kreisler (a conserved Maf b-zip protein) is required for the formation of the r4 posterior boundary and regulates the expression of *Hoxb3* and *Hoxa3* in r5-r6 (Cordes and Barsh, 1994; Moens *et al* 1996; Moens *et al.*, 1998; Manzanares *et al.*, 1999a; Manzanares *et al.*, 1999b,). In *Kr-/-* mice embryos the neural tube posterior to r4 appears unsegmented, a defect that is attributable to the loss of r5 and r6 as identifiable territories. However, both *Krox20* and *Kreisler* expression in prerhombomeric territories is transient, while segmental expression of anterior *Hox* genes persists for a longer period (Nieto *et al.*, 1991; Cordes and Barsh, 1994; Moens *et al.*, 1996; Manzanares *et al.*, 1999a). This observation can be explained by two different mechanisms. In one model, the vertebrate Polycomb and Trithorax protein complexes are able to lock the chromatin structure in such a way that maintains the proper *Hox* expression pattern established in early embyogenesis is maintained (Paro *et al.*, 1990 and 1993).

The Polycomb group (Pc-G) interacts with many genes in *Drosophila* as well in vertebrate, but the best understood role is in the regulation of homeotic genes. Pc-G members form multiprotein complexes establishing a silenced state

an specific genes (Pirrotta, 1997). Targeted disruption of *rae28*, a homolog of the *Drosophila* polyhomeotic gene, shifts the expression of *Hoxb3* and *Hoxd4* by one rhombomere in the anterior direction (Takihara *et al.*, 1997).

Other chromatin modulators belong to the Trithorax group, which are zinc-finger containing proteins that act as positive regulators, binding active chromatin and modulating *Hox* expression by counterbalancing the Pc-G effect. Additional chromatin modifiers that maintain an open chromatin conformation are the Brahma ortologues (Brg1- Snf2/Swi2).

The second model involves auto-, cross and para regulatory loops between *Hox* genes, whereby they themselves reinforce expression triggered by an independent process (Fig 1.12). Transgenic mice analysis has shown that the *Hox* and their cofactors play a fundamental role in these maintenance mechanisms. For example Hox-Pbx protein complexes (Fig 1.12) are important for maintaining the segmental expression of *Hoxb1*, *Hoxb2*, *Hoxb4* and *Hoxa4* (Popperl *et al.*, 1995; Maconochie *et al.*, 1997; Gould *et al.*, 1997; Studer *et al* 1998; Gould *et al.*, 1998) (this topic will be discussed in detail below).

In conclusion, a model that could explain the specific expression of *Hox* genes involves firstly the morphogens RA and FGF (Fig 1.12). In the hindbrain the cascade that activates the 3′ *Hox* genes starts by establishing an active RA gradient/distribution that results in the activation of *Hoxa1*. Hoxa1 and RA are then involved in establishing the Hoxb1 expression pattern along the hindbrain. In the later stages, *Hoxb1* is required to maintain its own synthesis and r4 identity (Studer *et al.*, 1996; Studer *et al.*, 1998). (The enhancer organization of *Hoxb1* will be analyzed in detail below). Many experimental observations suggest that Hox-1 proteins, together with Pbx, function at the top of the hindbrain patterning hierarchy, possibly specifying rhombomere identities. Actually, Pbx proteins are ubiquitously distributed along the hindbrain and they act not only on the Hox-1 group, but also on other *Hox* genes to specify aspects of r2-r6 identity.

How this gradual sharpening of expression boundaries occurs, and how these boundaries are subsequently maintained, are important questions in understanding the regionalization of the hindbrain. Cells of adjacent rhombomeres show different adhesive properties that prevent them from mixing, and this suggests a mechanism for hindbrain regionalization. Ephrins, membrane-bound proteins that interact with cell surface receptors (Eph) on adjacent cells, seem to be involved in this process which can generate bidirectional signals. Supporting this model, it has been shown that several Ephs and Ephrin genes are expressed in complementary rhombomere-restricted patterns in vertebrate embryos (Xu *et al.*, 1999). Thus, cells move across boundaries to join the appropriate territory, driven by the action of Ephrins/Ephs families of ligands and receptors, which mediate cell sorting throughout the rhombomeres. Therefore, the normal signaling between adjacent Eph or Ephrin expressing cells at the rhombomere borders leads to segregation of cells.

This idea is supported by the observation that cells from odd or even-numbered rhombomers have opposite adhesive-repulsive properties. Cells from the even ones are able to mix with cells from other even-numbered rhombomeres, while cells from odd mix with other odd numbered rhombomers. By contrast, cells taken from odd and even-numbered rhombomeres when mixed

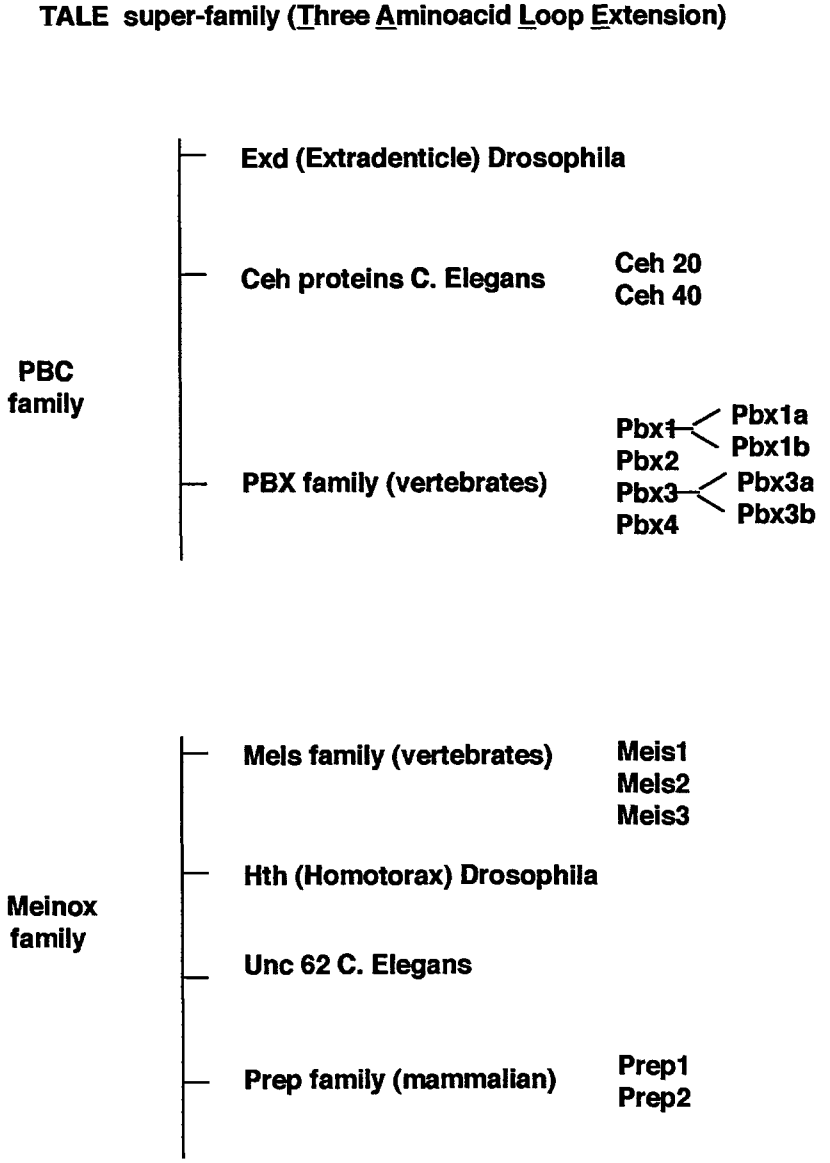


Figure 1.13
Schematic representation of TALE superfamily members

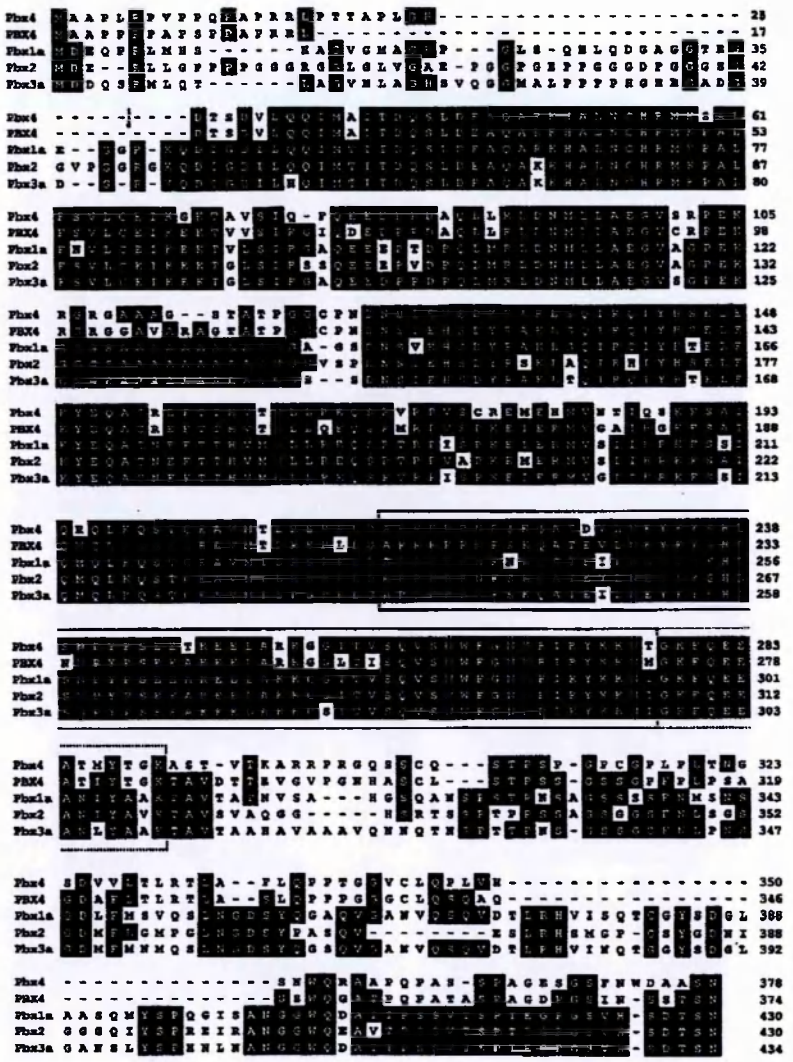


Figure 1.14

Alignment of aminoacid sequences of murine
Pbx1, Pbx2, Pbx3, Pbx4 and PBX4. The conserved domains domains PBC-A, PBC-B are boxed in black and conserved homeodomain (HD) underlined by solid lines (figure from Wagner *et al.*, 2002).

together segregate in cultured aggregates *in vitro* (Wizenmamm and Lumsden, 1997).

Cell transposition experiments in mouse have shown a degree of plasticity with respect to *Hox* expression not only for cranial neural crest cells but also for the rhombomeric tissues (Trainor and Krumlauf, 2000). This plasticity indicates that once the boundaries are formed, cells dispersed in an inappropriate location may be able to change their fate The nature of the signaling process that drives cells to change their identity remains to be determined (Trainor and Krumlauf, 2000). However, it seems clear that establishment and maintenance of rhombomere boundaries is likely to be a complex process, involving both cell sorting and plasticity.

DNA-binding specificity of Hox

Although each Hox protein exhibits *in vivo* specificity, as demonstrated by both loss and gain of function experiments, *in vitro* they all bind the same target sequences with low affinities. This paradoxical observation raises the question of how multiple proteins with similar DNA-binding specificity can activate or repress different, multiple target genes *in vivo*. Several studies show that homeodomain proteins rely on protein-protein interactions to confer enhanced DNA binding affinity and specificity or enhanced transcriptional activity. The vertebrate Pbx and Meinox proteins, encoding divergent homeodomain-polypeptide proteins of the TALE family (Fig 1.13), are essential DNA binding partners for the Hox proteins, specifying their DNA-binding ability and selectivity (Mann and Afflolter, 1998).

PBC proteins

The first candidates proposed as Hox cofactors were the homeodomain containing proteins PBC, including the vertebrate Pbx, the *Drosophila* Extradenticle (Exd) and the *C Elegans* Ceh-20 and Ceh-40 (Fig 1.13; Fig 1.14) (Burling *et al.*, 1995).

The PBX genes were originally identified upon cloning of the product chromosomal translocation (1: 19), fusing part of the bHLH domain of the transcription factor gene *E2A* to the homeodomain of PBX1. This translocation is present in 25% of cases of pre-B-cell acute lymphoblastoid leukemia (pre-B-ALL), hence the name pre-B cell homeobox (PBX) (Kamps *et al.*, 1990; Nourse *et al.*, 1990; Kamps *et al.*, 1991). While the E2A-PBX1 protein shows a strong transcriptional activation property, activating several genes in different cell types, the PBX proteins alone do not have this activity (Feldman *et al.*, 1997; Lu and Kamps, 1997; McWhirter *et al.*, 1997; Lu *et al.*, 1999a/b; McWhirter *et al.*, 1999).

The oncogenic fusion protein E2A-PBX, but not PBX1 alone, can activate the transcription of reporter genes containing the Pbx consensus binding sites (PRS: TGATGAT), suggesting that PBX1 does not have a transcriptional activation domain (Lu *et al.*, 1994; van Dijk *et al.*, 1993).

The Pbx family includes four genes Pbx1, Pbx2, Pbx3 and the recently cloned Pbx4 (Fig 1.13 and Fig 1.14; Monica *et al.*, 1991; Popperl *et al* 2000). Pbx1 and Pbx3 have two alternative splice variants (Pbx1a-1b, Pbx3a-3b) with different C-terminal regions.

The Pbx proteins show a high degree of homology, although Pbx4 seems to be the most divergent gene. The homology between Pbx proteins extends beyond the homeodomain and includes two other regions located at the N-terminus, called the PBC-A and PBC-B domains. PBC-A and PBC-B are important protein-protein interaction domains mediating dimerization with Meinox proteins (see below).

Expression and function of the PBC protein family.

The Pbx proteins are expressed in different tissues in adult and fetal mice. They also show different expression levels in different tissues, and they have organ specificity. Different levels of expression of Pbx are also observed during embryonic mouse development. Pbx1 is strongly expressed until E 10.5 d.p.c, but later its synthesis is reduced (Ferretti *et al.*, 1999). Pbx3 is highly expressed in the developing central and peripheral nervous system, particularly in the medulla oblongata (Torreson *et al.*, 2000). In the adult, Pbx1 is ubiquitously present, with the exception of lymphoid tissues (Monica et al., 1994). Pbx2 and Pbx3 are also ubiquitously expressed in adult murine tissues. However, Pbx2 is not found in the heart, while Pbx3 is predominantly present in adult ovary (Monica *et al.*, 1994).

Pbx4 is highly expressed in mouse testis and is not present in mouse embryos between 7 and 10.5 d.p.c. (Wagner *et al.*, 2001).

Genetic inactivation of Pbx1 in mice has shown that Pbx1 is an essential gene, and its loss results in late gestational lethality (embryos die in uterus at 15.5 d.p.c.) accompanied by severe hypoplasia (of lungs, liver, stomach, gut, kidney and pancreas), the presence of ectopic thymus and kidneys, and finally the aplasia of spleen (Selleri *et al.*, 2002). The role of Pbx1 in segmentation is shown by the diffuse defects of *Pbx1-/-* embryos in patterning of the axial and proximal appendicular skeleton, like its homolog Exd in *Drosophila* (Rauskolb et al., 1995; Selleri et al., 2002). Loss of *Pbx1* is accompanied by a homeotic-like transformation of the second branchial arch (BA2) to the first branchial arch (BA1), similar to the phenotype shown in the zebrafish *lazarus* mutant in which the Pbx4 gene is inactivated (Popperl *et al.*, 2000; Selleri *et al.*, 2002).

Furthermore, absence of Pbx1 causes abnormalities in chondrocyte proliferation and differentiation, which results in precocious ossification and bone formation (Selleri *et al.*, 2002). The effect on skeletal malformations observed in Pbx1-/- mice does not precisely recapitulate individual *Hox* gene inactivation, suggesting a possible *Hox*-independent role for Pbx1 (Selleri *et al.*, 2002). In addition, it must be taken into account that different Pbx proteins can interact with different Hox proteins (see below).

While the *Pbx2* knockout mice do not present any visible phenotype (Selleri, personal communication), *Pbx3-/-* mice die a few hours after birth due to central respiratory failure. This death is dependent on the abnormal activity of inspiratory neurons in the ventrolateral medulla (Selleri, personal communication).

Furthermore, in zebrafish embryo, inactivation of *lazarus* (homologous to mammalian *Pbx4*) causes alteration in A/P identity through the hindbrain and anterior trunk. *Lazarus* mutants exhibit the absence of rhombomere boundaries anterior to r4, the fusion of the first and second pharyngeal arch cartilages (characteristic of the *Hoxa2* mutant) (Rijli *et al.*, 1998; Popperl *et al.*, 2000). Moreover, *lazarus* mutants show anterior homeotic transformation of the primary reticulo-spinal neurons and the facial neurons migrate posteriorly from r4 (characteristic of *Hoxb1* -/- mice) (Studer *et al.*, 1996; Popperl *et al.*, 2000). Furthermore, the inactivation of both Pbx4 and Pbx2 leads to the total loss of hindbrain segmentation, revealing a background r1 state (Waskiewicz., et al 2002).

Pbx-Hox heterodimerization.

The studies on Pbx inactivation demonstrated an important role for the Pbx protein in processes that are the territory of the various Hox genes. Indeed, multiple studies have described Pbx-Hox heterodimerization and shown that Hox-Pbx heterodimers bind the specific bipartite sequence 5'TGATNNAT [g/t] [g/a]3' (Mann, 1996; Di Rocco *et al.*, 1997). By methylation interference, it has been demonstrated that in the Pbx-Hox heterodimer the Pbx protein binds the TTGAT part of the sequence, while Hox contacts the sequence NNAT. The two base pairs NN are predicted to contact the N-terminal arm of the Hox homeodomain and seem to have a crucial role in selecting which Hox partner is preferred in dimerization (Chang *et al.*, 1996; Knoepfler *et al.*, 1996; Chan *et al.*, 1997; Manzanares *et al.*, 2001).

In addition, co-immuno-precipitation experiments have suggested that Hox-Pbx forms stable heterodimers only in the presence of specific target consensus sites.

Several observations have demonstrated that the Hox-Pbx interaction requires two domains of Hox, the homeodomain and an N-terminal conserved hexapeptide, and two structures of Pbx, the homeodomain and a conserved C-terminal 15 amino acid stretch. The "hexapeptide" sequence (IYPWMK), or its core motif (YPWM), is required for the cooperative DNA binding of Pbx and Hox proteins to consensus binding sites (Chang *et al.*, 1995; Phelan *et al.*, 1995). The YPWM motif is found at a variable distance (5-50 amino acid) N-terminal to the homeodomain in the Hox paralogues 1 to 8. The tryptophan residues that mediate cooperative binding are also found in paralogs 9 and 10 and in other non-clustered *Hox* genes.

Interestingly, the hexapeptide appears to have co-evolved with its associated homeodomain, suggesting that these two domains may function together. In addition to being an important component of the PBC-Hox interaction, the hexapeptide might also act as an inhibitor of DNA binding, as demonstrated for Labial. Other Hox amino acids, perhaps within or C-terminal to the Hox homeodomain, are also necessary for the interaction with PBC proteins.

In addition to the homeodomain, the presence of the unusual three amino acid loop of Pbx appears to be important, suggesting that this loop might constitute part of the Hox interaction surface. Moreover, a stretch of 15 amino

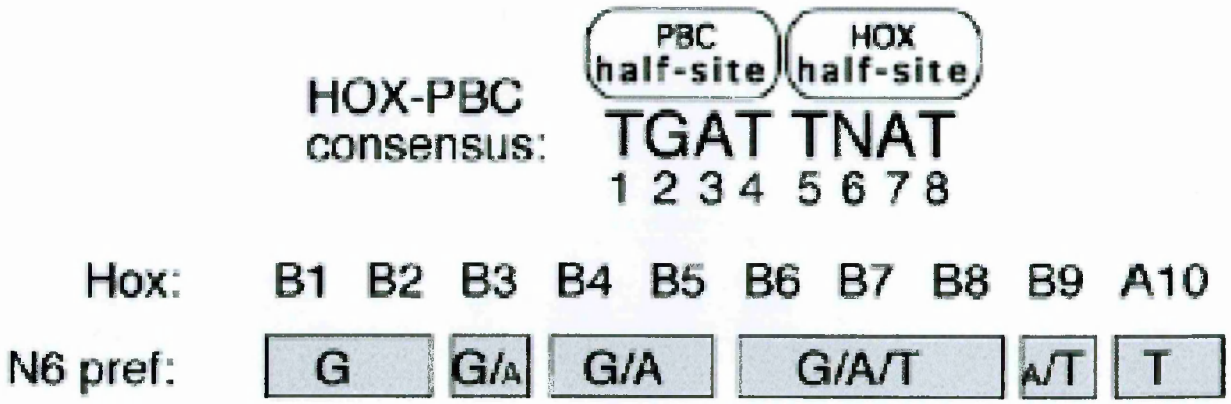


Figure 1.15
The Hox-PBC binding site is recognized by different dimers. Yhe binding specificity of the various PBC-Hox complex is dependent on the sixth position of the consensus sequence of the target site (from Wilson *et al.*, 1999).

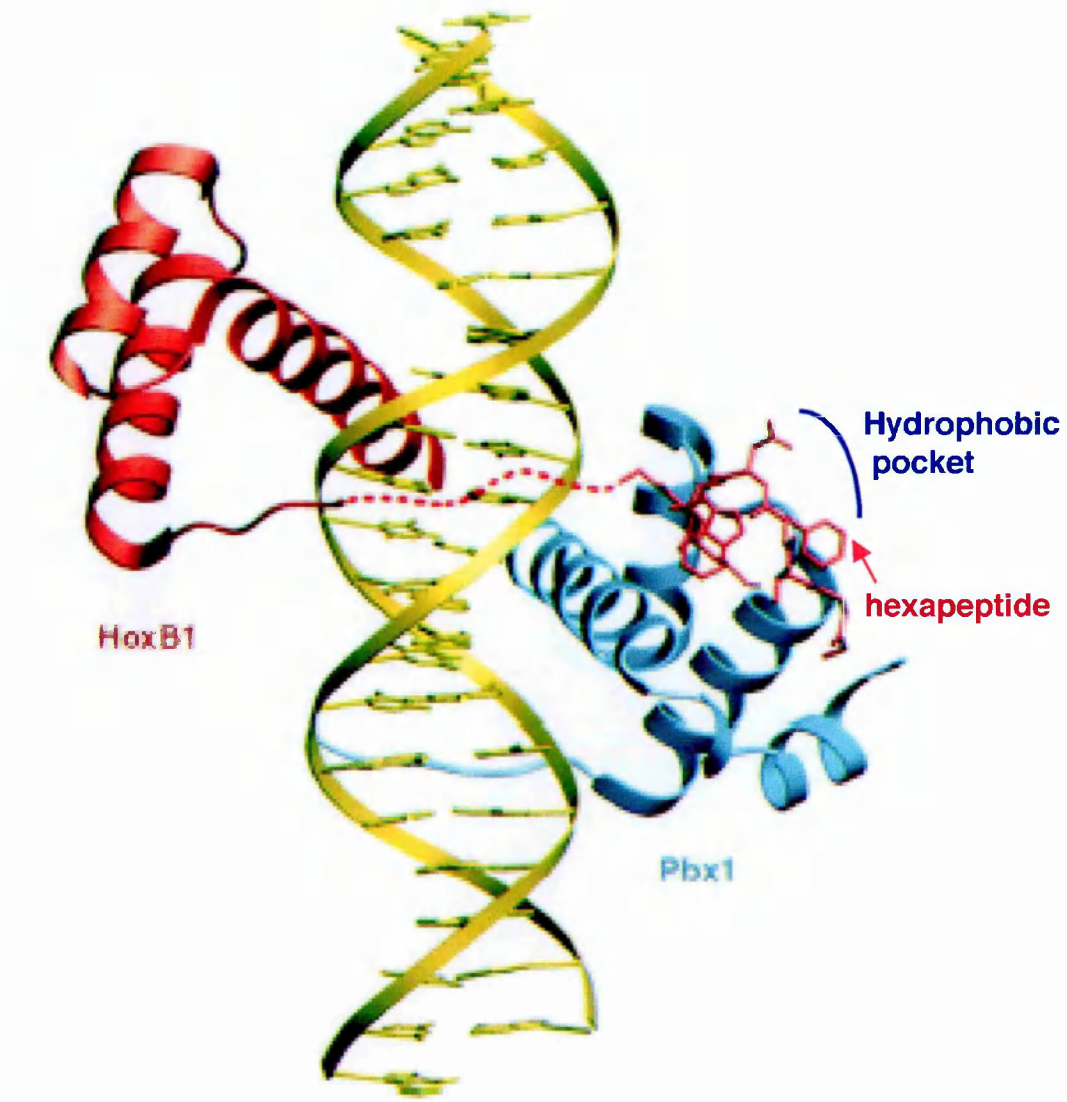


Figure 1.1.6

Ribbon diagram of the Hoxb1 and Pbx1 homeodomain bound to DNA. HOxb1 hexapeptide, which is connected to the N-Terminal of the Hoxb1 homeodomain by flexible linker, contacts the Pbx DNA-binding domain. The C-termibnal 15aa peptide of Pbx1 forms a helix stabilizing helixIII of Pbx1 and not contacting DNA

acids C-terminal to the PBC homeodomain is also required for the Pbx-Hox interaction.

The Hox proteins belonging to the parologue groups 1 to 10 are capable of binding the PBC proteins. On the contrary, the posterior Hox proteins, paralogue groups 11-13, are not able to interact with the PBC proteins, and are missing the hexapeptide. However, other TALE proteins, for example Meis, are able to directly interact with these Hox proteins (see below).

Hox-Pbx binding specificity

Although most of the Hox-Pbx complexes bind the TGATNNAT consensus-sequence, binding site selection experiments performed with Pbx1 and Hox-B paralogous proteins have proposed the TGATTNAT sequence as a specific *PBC/Hox* target site (Fig 1.15) (Chang *et al.*, 1996). These studies have shown that the DNA binding selectivity of the Pbx1-Hox-B heterodimer depends on the nature of the nucleotide (N) in position 6. In fact, the most anterior Hox proteins (Hoxb1 and Hoxb2) prefer G, while Hoxb3, Hoxb4 and Hoxb5 prefer G or A. The middle Hox-B paralogue (Hoxb6-Hoxb7) prefer G/A/T, while the most posterior Hox (Hoxb9 and Hoxa10) choose T only. Thus the nature of the nucleotide in position 6 is crucial in defining which Pbx-Hox heterodimers will bind, suggesting that the 3′ part of the bipartite consensus is conferring the DNA binding specificity (Fig 1.15).

Structure of the Pbx1-Hoxb1 heterodimer

Recently, the crystal structure of the Hoxb1-Pbx1-DNA complex has been defined (Pipper *et al.*, 1999; Passner *et al.*, 1999; Sprules *et al* 2002). Structural data show that helix III of the two homeodomains contacts the major groove of the DNA (see Fig 1.16). The two homeodomains are located on opposite sides of the DNA in a head to tail orientation, resulting in a separation of almost a half turn of the DNA helix.

Thus, the two homeodomains do not interact with each other, and the connection between the two proteins is constituted by the Hox hexapeptide that is able to contact the three aminoacids loop (characteristic of the TALE class of proteins) through a flexible linker, forming a hydrophobic pocket. Recently, a model of key and lock has been proposed, in which the insertion of the Hox hexapeptide key into the hydrophobic lock on the surface of the Pbx homeodomain tethers the Hox homeodomain to the DNA duplex, resulting in the formation of more stable contacts between the Hox N-terminal arm and DNA. Unexpectedly, the 15 amino acids C-terminal to the Pbx homeodomain, which are important for the DNA binding properties of the dimers, do not contact the DNA, but form a helical extension probably stabilizing helix III-DNA interaction.

Modulators of Pbx-Hox activity

Although the binding selection role for the PBC proteins has significant experimental support, it cannot explain completely the Hox specification observed *in vivo*.

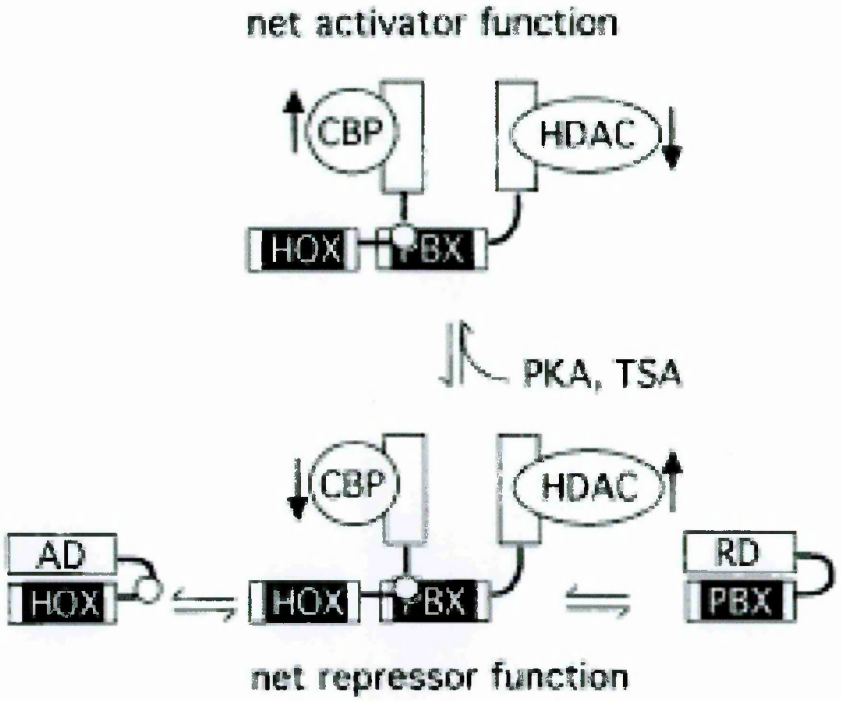
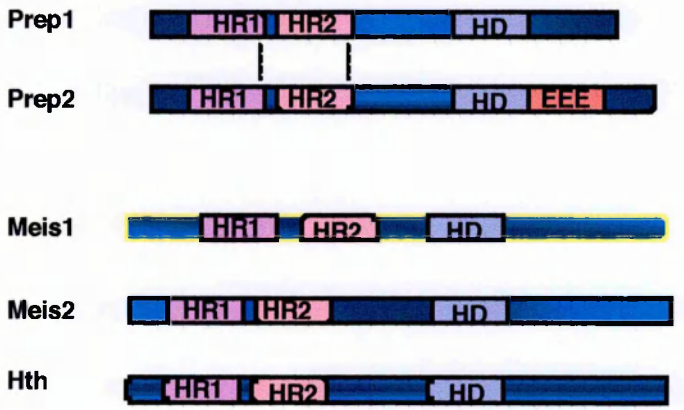


Figure 1.17

Model for activation and repression by Hox-Pbx complexes. The N-terminal activation and repression domains of Hox and Pbx proteins, respectively, are believed to make intramolecular contacts with their respective homeodomains. Heterodimerization on cooperative sites on DNA and perhaps additional interactions with members of Meis - Prep1 family exposes or hides Pbx N-termini, thereby freeing them for interaction with co-activators and co-repressors such as CBP and HDAC1 and 3. However, PKA signaling increases co-activator and/or decreases co-repressor function, shifting the balance toward net activation. Treatment with TSA would exert the same overall effect by inhibiting bound HDACs (from Saleh *et al.*, 2000).

The Meinox sub-family




```
Prep  DADKQAIYRHPLFPLLALLFEKCEQST                KGRVLPKHATNVMFSWLFQHTGHPYPTEDEKKQIAAQITN LTLLQVNNWFINNARILQPMLD
Meis  KRDKDAIYGHPLFPLLALI FEKCELAT                KGRI FPKVATNI MFAWLFQH LTHPYPSEEEKKQLAQDITGLTI LQVNNWFINNARIVQPMI D

Prep DPETDNLMYKAIQYLRI HLLGLEKYNELCKDFCSRV I ACLKTKM
Meis DPETDNLMYQAIQYLRFHLLGLEKYNELCDNFCHRVI SCLKGKM
```

Figure 1.18

Structural alignment of Prep1, Prep2, Meis1 , Meis 2 and Hth proteins showing the three conserved domains HR1 and HR2 (Homology Region 1 and 2) and HD (homeodomain). Below is shown the sequence alignment between murine Prep1 and Meis1 of the Meinox interaction domains (HR1 and HR2) and HD. All Meinox proteins contain these homology regions.

Several studies have proposed that the Hox proteins can act as activators or repressors through differential association with co-activators and co-repressors in a context-dependent manner (Struhl *et al.*, 1998). By analogy, Hox-Pbx complexes may do likewise to achieve transcriptional repression or activation. Pinsonneault *et al.*, suggested that Exd and Pbx could switch the Hox proteins from repressors to activators (Pinsonneault *et al.*, 1997). The switch of the activity status seems to depend on the presence of coregulators such as the histone acetyltransferase (HATs) and the histone deacetylases (HDACs), which modify chromatin, as well as non histone proteins. A member of the HAT family (CBP/p300) has been found in association with Hox proteins (Shen *et al.*, 2001) and with Hox-Pbx dimers (Chariot *et al.*, 1999). CBP/p300 proteins are known to enhance the transcription potential of a variety of proteins and to act as a bridge or an adaptor between the DNA-binding regulators and the basal transcription machinery (Shen *et al.*, 2001). There are two possible mechanisms by which these enzymes can act. One is the acetylation process, whereby CBP/p300 presumably destabilizes the nucleosome and facilitates the access of transcription factors to DNA. An alternative is that CBP and p300 directly acetylate transcription factors thereby altering their DNA binding capacities

It has been reported that CBP/p300 specifically interacts with Pbx1-Hoxb7 heterodimers, contacting the N-terminus of Hoxb7 and resulting in the enhancement of the transcriptional activity (Saleh *et al.*, 2000). The observation that the protein kinase A (PKA) pathway significantly potentiates CBP-mediated transactivation by Hox-Pbx complexes leads to a model in which PKA and HDAC act as signaling switch factors that convert the Hox-Pbx complex from a transcriptional repressor to an activator (Fig 1.17, Saleh *et al.*, 2000).

Using yeast two hybrid screening of a fetal liver hematopoietic cDNA library, a novel non-homeodomain protein: Hematopoietic PBX Interacting Protein (HPIP) has been discovered that is able to interact with different members of the Pbx family. HPIP is mainly localized in the cytosol and in small amount in the nucleus of early embryonic and adult murine tissues. The region of Pbx that interacts with HPIP includes both the homeodomain and immediate N-terminal flanking sequences. EMSA reveals that the interaction of HPIP with Pbx inhibits Pbx-Hox DNA-binding activity, suggesting that HPIP could represent another step in Pbx-Hox regulation.
HPIP contains both nuclear import and export sequences, but it seems to be constitutively associated with cytoskeletal structures, suggesting a model in which after unknown stimuli the molecule shuttles between the nucleus and the cytosol, thus regulating Pbx function (Abramovich *et al.*, 2002).

The Meinox group

The other classes of Hox partners belonging to the TALE super-family are the Meinox proteins (Fig 1.13), including the vertebrate Meis and Prep families, which are tightly regulated and appear to have important Pbx-dependent function during embryonic development. These proteins interact with Pbx even in the absence of DNA. The sequence similarities within the Meinox sub-family are essentially confined to the homeodomain and to the two N-terminal regions,

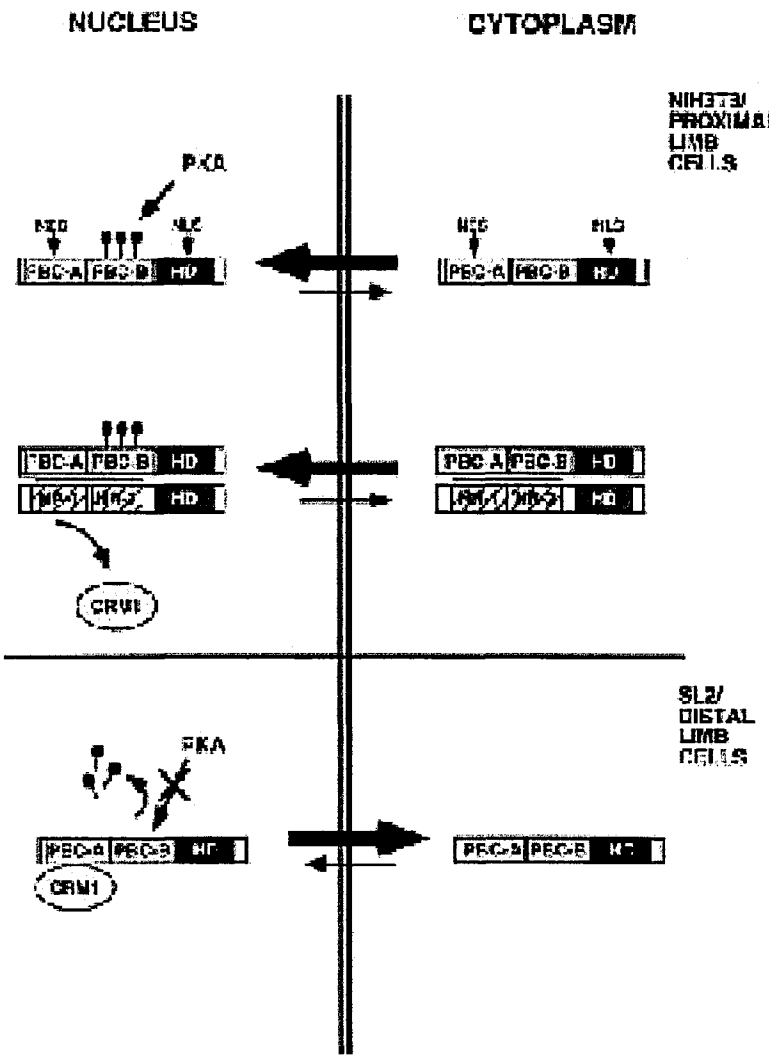


Figure 1.19

PKA activity controls the subcellular localization of Pbx proteins. Pbx1 nuclear export, which requires a signal located within the PBC-a domain, is prevented by heterodimerization with the Meinox protein (Berthelsen *et al.*, 1999). Hence, nuclear export was assumed to be the default route for subcellular localization of PBC proteins in the absence of Meinox proteins. Pbx1 was found to contain two adjacent independent nuclear export signals of the leucine-rich type, located within the conserved PBC-A domain, which mediate direct contact with the CRM1 export receptor. The cellular distribution is controlled by the PBC-B domain and correlates with the phosphorylation of this serine residue *in vitro* and PKA activation blocks nuclear export of Pbx1 in distal regions of the developing limb bud *in vivo* (from Kilstrup-Nielsen e al 2003).

which are involved in protein-protein interactions with PBC proteins, and are termed MH (Meis-Homology) or HR (Homology Region) (Fig 1.18). Meinox-Pbx interaction does not require the respective homeodomain and in addition, complex formation precedes DNA binding as deletion of the N-terminal part of either protein disrupts co-operative DNA binding of the PBC-Meinox complexes. However, in order to bind the DNA, the heterodimers require the presence of both homeodomains. The use of different interaction surfaces allows PBC proteins to act as a bridge between Meinox and Hox proteins to form transcriptionally active ternary complexes (see below).

Pbx-Meinox heterodimers show the same *in vitro* binding specificities. They bind to TGACAG or TGATTGACAG sequences, sites not recognized by Pbx-Hox dimers, but in addition they can interact with the Pbx-Hox target site (TGATNNATNN) (Chang *et al.*, 1997; Koefpler *et al.*, 1997; Berthelsen *et al.*, 1998; Jacobs *et al.*, 1999: Ferretti *et al.*, 2000). Moreover, unlike Pbx-Hox complexes, the Pbx-Meinox dimers do not show direct transcriptional activation properties (De Cesare *et al.*, 1996; Goudet *et al.*, 1999; Berthelsen *et al.*, 1998) leaving open the question of the function of these heterodimers. The Meinox proteins are subdivided into two families, Prep and Meis (see below), on the basis of their structural and expression domain similarities.

Regulation and function of PBC-Meinox dimers

Several intriguing aspects of the function of PBC and Prep/Meis were discovered during genetic studies of the *Drosophila* orthologs *Exd* and *Hth*, respectively.

As mentioned previously, Exd/Pbx activity is regulated postranscriptionally by the nuclear translocation regulatory mechanism.
In *Drosophila*, Exd is essential to establish the distal-proximal axis in the developing limb-bud. Exd is required for patterning of the proximal region of the fly legs, where Exd is subjected to nuclear-cytoplasmic regulation. In fact, while Exd is expressed uniformly in leg imaginal discs, its subcellular localization is regulated. In the proximal part of the limb bud (developing legs), Exd is localized in the nucleus and its activity prevents the cells from responding to decapentaplegic (Dpp) and Wingless (Wg). On the contrary, in the distal region of the limb bud, Exd is cytoplasmic, thus in an inactive form, allowing the activation of Dpp and Wg signaling proteins, which are in turn activated by Hh (Hedgehog) (Consalez-Crespo *et al.*, 1998). The positive regulator of Exd nuclear translocation is the Meinox protein Hth. In the absence of Hth, Exd remains in the cytoplasm and it is not functional. In the presence of Hth, Exd translocates to the nucleus, a process that requires direct interaction between Exd and Hth (Rieckhof *et al.*, 1997; Kurant *et al.*, 1998; Pai *et al.*, 1998). Thus, both Hth and Exd are required for proximal, but not distal, leg development. Ectopic expression of Hth distally induces nuclear localization of Exd, blocking the distal development and resulting in truncated appendices (Morata, 2001).

The correlation between proximal distal limb bud patterning and subcellular localization is evolutionarily conserved (Consalez-Crespo *et al.*, 1998; Capdevilla *et al.*, 1999; Mercander *et al.*, 1999). Furthermore, Meis1 and Prep1 can also substitute for Hth to induce the nuclear localization of Exd in salivary gland and antennal and leg imaginal discs (Rieckhof *et al.*, 1997; Berthelsen *et al.*, 1999;

Jaw *et al.*, 2000). Misexpression of Meis1 or Meis2 in distal chicken limb buds produces proximalization or truncation of distal structures (Capdevilla *et al.*, 1999; Mercander *et al.*, 1999). Ectopic expression of human Prep1 in Drosophila imaginal discs weakly affects eye development but strongly affect leg and antennal development (Jaw *et al.*, 2000).

Likewise, in the developing mouse limb bud, Pbx1 is cytoplasmic in the distal part and it is nuclear only in the proximal part, due to the presence of the Hth orthologs Meis 1 and 2 (Fig 1.19, Consalez-Crespo *et al.*, 1998, Mercander *et al.*, 1999; Kilstrup-Nielsen *et al.*, 2003). Both Exd and Pbx proteins contain a nuclear localization signal (NLS), as well as a nuclear export signal (NES). The NES signal is mapped in the PBC domain, which is also involved in Meinox protein interaction. Since Meinox does not carry an already defined (NLS), a model has been proposed in which the heterodimerization of Exd-Pbx with Hth-Meinox induces nuclear translocation of Meinox-Pbx dimers, which is then kept in the nucleus through the blocking of the NES (Fig 1.19) (Bethelsen *et al.*, 1999). Recently two different mechanisms for controlling Pbx1 nuclear export, in which the Meinox export is not involved, have been reported. One involves the interaction of two leucine rich sequences in the PBC-A domain with the CRM1 export receptor. The second involves the activity of Protein Kinase A (PKA). PKA mediates phosphorylation of specific serine residues within the PBC-B domain and thus blocks the nuclear export of Pbx1 (Fig 1.19) (Kilstrup-Nielsen *et al.*, 2003).

Pbx/Meinox balance mediates stabilization.

The regulation of the sub-cellular localization of the PBC/Meinox complexes suggests that the control of the relative amounts of these two groups of proteins is important for *Hox* gene function. Recently another level of regulation of PBC/Meinox activity has been proposed, based on data obtained in Drosophila , Zebrafish and mouse studies. These studies showed a bi-directional stabilization between Meinox and PBC proteins. In Drosophila, Hth mutants exhibit a very low level of Exd protein with normal mRNA levels (Kurant et al., 1998). In Zebrafish, Meinox protein overexpression is able to regulate the level of Pbx proteins by a post-transcriptional mechanism, which requires protein-protein interaction. In fact, Meinox proteins which carry a deleted HM (or HR) domain, but no homeodomain, do not affect Pbx protein level (Waskiewicz *et al.*, 2001). The effect may be due to protein stabilization. These data suggest that the Meinox proteins may cooperate to stabilize Pbx/Hox complexes and thereby promote or affect *Hox* function during development. Moreover, Prep1 zebrafish morphants show a drastic reduction of all Pbx proteins in the embryo (Argenton and Blasi groups, unpublished data). As it will be shown in this thesis, *Prep1* knock out mice also display a great reduction in all Pbx proteins.

The Meis sub-family

The Meis protein family includes mammalian Meis1, Meis 2 and Meis 3, the *C. Elegans* unc 62, the *Drosophila* homotorax (Hth) and other orthologues in Zebrafish and *Xenopus Laevis* (Fig 1.13) (Burglin *et a.*, 1997; Steelman *et al.*, 1997).

Meis1 was discovered as a site of integration of the ecotropic murine myeloid leukemia virus that causes leukemia in mice, hence the name myeloid ecotropic integration site 1 (Meis1). It was observed that cooperativity of Meis1 and Hox proteins is required to induce leukemogenesis (Moskow *et al.*, 1995; Smith *et al.*, 1997). In addition to Pbx, Meis1 is able to directly bind Hox proteins of paralogue groups 9-13 through its C-terminal sequence. Moreover, Meis 1 binding of Hoxa9 induces transformation of primary bone marrow cells towards a myeloid leukemic phenotype (Kawagoe *et al.*, 1999; Lawrence *et al.*, 1999). Hoxa9 and Meis1 are down-regulated during normal myeloid differentiation. They are able to cooperate to induce acute myeloid leukemia (AML) in mice, and are co-expressed in human AML. The interaction of Hox with Meis proteins increases the binding specificity of Hox9-Hox13, suggesting a role for Meis1 in complementing the function of posterior Hox (Shen *et al.*, 1997).

Meis proteins seem to have a specific pattern of expression during development. Meis 2 is expressed in the developing nervous system, limbs, face and in various viscera (Cecconi *et al.*, 1997; Oulad-Abdelghani *et al.*, 1997). Meis1 and Meis 2 have a high level of expression in the mouse telencephalon, in particular in the subventricular zone (SVZ) and mantle regions of the ventral telencephalon (Toresson et al., 2000).

In Zebrafish, Meis proteins are expressed early in development in a region-specific pattern (Waskienwkz *et al.*, 2001; Biemar *et al.*, 2001; Choe *et al.*, 2002).

The Prep sub-family.

Prep is the last member of Meinox sub-family more recently identified (Fig 1.13). Prep and Meis family members show strong structural and functional similarities (Fig 1.18). Prep1 was identified as a major component of the PBC/Meinox dimers present in HeLa cells, and it binds the enhancer of the *urokinase* gene (Fig 1.20). Prep1 and Meis proteins are very similar in their homeodomains. In particular, they are almost identical in the third helix (the DNA recognition helix) and have very similar basic residues in the homeodomain N-terminal arm. In addition, they all have an isoleucine at residue 50 of the homeodomain, corresponding to the third position of the highly conserved homeodomain motif WF_N_N of the third helix. Position 50 is important in determining the DNA binding specificity, and isoleucine is rarely found at this position in the homeodomain. For all these reasons, the Meinox-Pbx dimers bind very similar sequences. Like Meis, Prep1 protein can associate with Pbx in the absence of DNA and bind the TGACAG sequence, as well as the *Pbx-Hox* consensus sites (Berthelsen *et al.*, 1998a; 1998b). Unlike Meis, Prep1 seems to be ubiquitously expressed both in adult and embryonic mouse tissues at least until 10.5 d.p.c. (Ferretti *et al.*, 1999). A new Prep family member, Prep2, has been recently discovered, showing structural and functional similarities with Prep1 (Fognani *et al.*, 2002; see below). This has allowed classifying Prep proteins as a sub-family of Meinox, distinct from Meis.

In addition to the anterior Hox and Pbx, Prep1 interacts with the pancreatic homeodomain transcription factor Pdx1 and modulates Pdx activation of the *somatostatin* gene (Goudet *et al.*, 1999) (Fig 1.20). Prep1-Pbx complexes have been also shown to repress the transcriptional activity of the *glucagon* gene in a

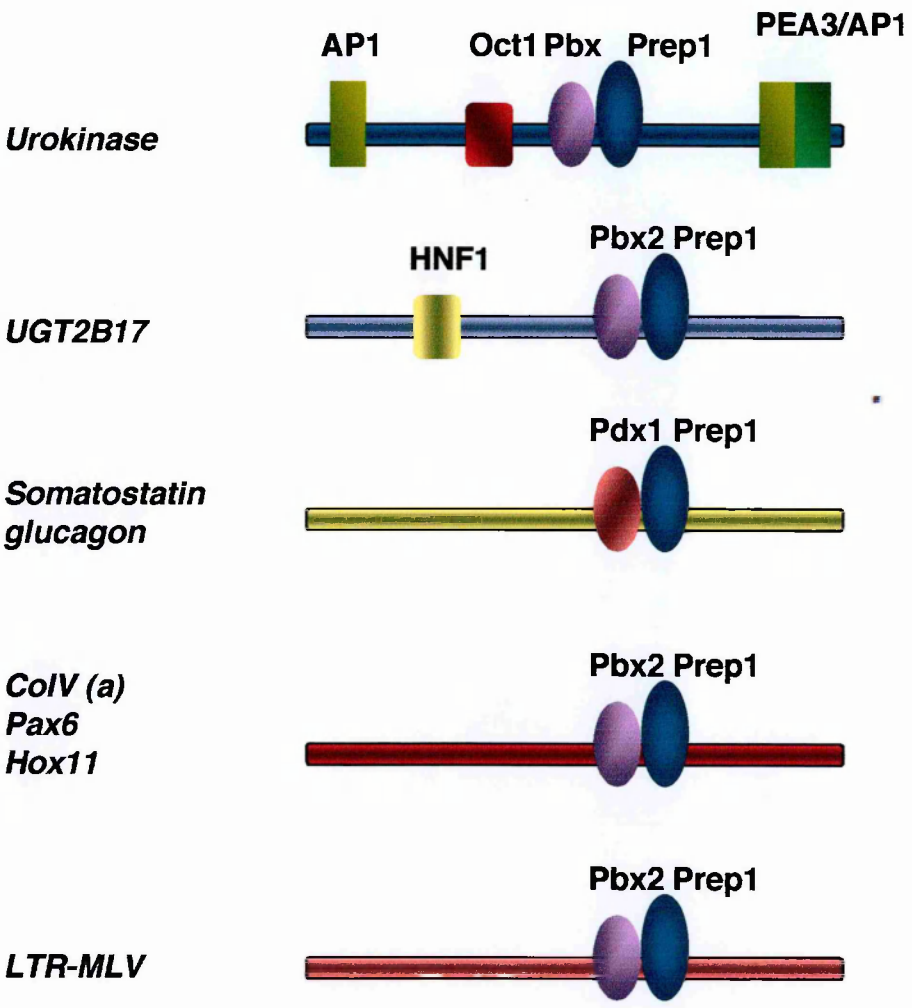


Figure 1.20
Schematic representation of Prep1-PBC heterodimers proteins binding to various enhancers elements (see text for details).

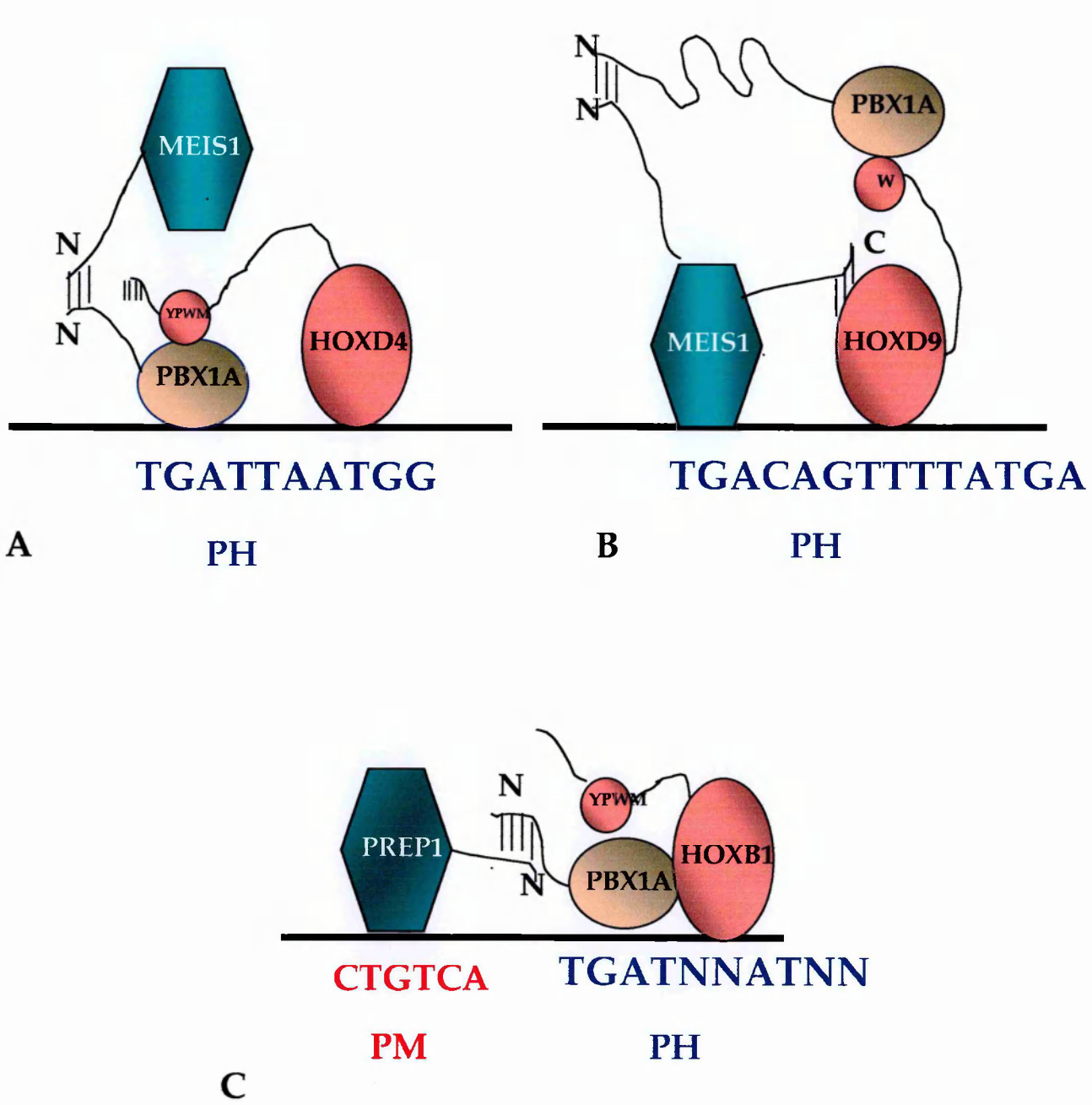


Figure 1.21

Model summarizing the formation of trimeric complexes among Hox, Pbx and Meinox proteins. A) Trimeric complex formed by Pbx1a, Hoxd4 and Meis1. Pbx and Hox bind the PH consensus, while Meis is the non DNA-binding partner. Meis and Pbx are interacting via N-termini. B) Trimeric complex formed by Pbx1a, Hoxd9 and Meis1. Pbx is the non DNA-binding partner and it is also interacting with Meis via the amino termini. C) Trimeric complex formed by Pbx1a, Hoxb1 and Prep1 which involves a combined PM and PH binding site, present in the *Hoxb1* and *Hoxb2* enhancers. The binding domains of all proteins are participating in the binding. Modified from Shanmugam *et al.*, 1999).

cell-type dependent manner (Herzig *et al.*, 2000) (Fig 1.20). Likewise, Prep1-Pbx2 complexes appear to be involved in the regulation of the *ColV(a)*, *Pax6* and *Hox11* genes (Fig 1.20) (Penkov *et al.*, 2000; Mikkola *et al.*, 2002; Brake *et al.*, 2002). Likewise, Pbx2-Prep1 complexes can repress the activity of the *UDP-Glucoronyltransferase 2B17* gene, modulating the activity of the transcriptional factor HNF1 (Fig 1.20) (Gregory and Mackenzie, 2002). Recently, it has been reported that Prep1-Pbx1 dimers are able to enhance the transcription of Murine Leukemia Virus (MLT) (Chao et al., 2003). Thus, taken together, the data suggest that Prep/Pbx dimers can act as activators or repressors in an enhancer specific manner. However, no information is as yet available on the role of the Prep protein in development. This thesis gives some clues towards answering this question.

Hox-PBC-Meinox ternary complex

Based on the functional and structural data described, it appears that evolution exerted a positive pressure on the conservation of PBC-Hox heterodimers. The stable heterodimer formation between PBC and Prep/Meis proteins is also well conserved. The observation that the domains involved in the interactions are different leaves the possibility that the three proteins may become physically linked forming a Hox-Pbx-Meinox heterotrimer.

Several heterotrimeric complexes have been identified resulting in three categories of ternary complexes, which differ in their TALE DNA-binding properties.

The first group includes complexes in which the Meinox proteins are not required to directly bind DNA. These complexes include ternary complexes formed by Pbx1a-Meis1-Hoxd4 that bind to the Hox-Pbx consensus (Fig 1.21 panel A), and the complexes containing the pancreas specific homeodomain protein Pdx (Pdx1-Meis2-Pbx1b and Pdx-Pbx1-Prep1) which bind to the *elastase I* and *somatostatin* enhancers, respectively (Swift *et al.*, 1998; Goudet *et al.*, 1999; Shamungam *et al.*, 1999). However, no real proof exists that these first type of complexes actually have a functional role *in vivo*.

The second category includes ternary complexes in which Pbx is the non DNA-binding partner in trimeric association, for example Meis-Hoxd9-Pbx or Meis-Hoxd10-Pbx are able to bind the PBC-Hox consensus (Fig 1.21, panel B).
The third are the heterotrimeric complexes (Hoxb1-Pbx1-Prep1 and Hoxb1-Pbx1-Meis1) in which all the binding domains of the proteins involved are participating in the binding to DNA.

In this thesis I also demonstrate that the ternary complex Hoxb1-Pbx1-Prep1 forms on *Hoxb1* and *Hoxb2* r4 responsive elements *in vitro*. These complexes require a combined Pbx-Hoxb1 (PH) and Prep-Pbx (PM) binding site, and they also have a functional role *in vivo* (Fig 1.21 C; Jacobs *et al.*, 1999; Ferretti *et al.*, 2000) (see chapter 4). A similar situation is described in Drosophila, where a Labial-Exd-Hth ternary complex activates the *labial* promoter, requiring DNA-binding of all three proteins (Ryoo *et al.*, 1999). Thus, this kind of DNA-binding ternary complex seems to activate the transcription of specific genes, and it is also conserved during the evolution of the fly to mammals.

CHAPTER 2

Aim of this work

The aim of my thesis was to characterize the role of the transcription factor Pbx, Regulating Protein1 Prep1 *in vivo* and *in vitro*. In order to address this, I first cloned the murine Prep1 cDNA and analyzed the expression of *Prep1* in adult and embryonic murine tissues. Subsequently, since it was known that Prep1 belongs to the TALE class of proteins and that it was able to form heterodimers with Pbx, I analyzed the presence of the Prep1-Pbx binding activity both during mouse embryogenesis and in adult murine tissues. The resulting data suggested that Prep1 represented a constitutive partner for Pbx proteins both in adult and in embryo. Since Pbx proteins are co-factors of *Hox genes* that regulate Hox-DNA binding activity and specificity, and it is known that Hoxb1-Pbx1 partecipate on its own expression in rhombomere4, I decided to address the function of Prep1 in Hox expression. This constitutes the main body of work presented in this thesis. Finally, I report the preliminary characterization of the Prep1 knock-out mouse.

In summary, my thesis describes the cloning of the murine Prep1 cDNA, the subsequent analysis of the function of Prep-Pbx complexes as modulators of Hox expression in the mouse hindbrain and in murine cultured cells and finally the preliminary characterization of a *Prep1* knock out mouse.

Specific background

PREP1 was originally purified from HeLa nuclear extracts as one of the polypeptides composing the Urokinase Enhancer Factor 3 (UEF3) complex binding to the *urokinase* enhancer (Fig 2.1) (Nerlov *et al.*, 1992; Berthelsen *et al.*, 1996). UEF3 from HeLa cells is able to bind the TGACAG sequence, forming two retarded bands in EMSA: UC (Upper Complex) and LC (Lower Complex) (see Fig 2.1). Subsequent analysis showed that both UC and LC were composed of a common 64 kDa polypeptide, PREP1, and of either a 50 kDa or 40 kDa subunit, corresponding to the human proteins PBX2 and PBX1B (Fig 2.1) (Berthelsen *et al.*, 1998a).

Protein micro-sequencing of p64-UEF3 provided a 17 amino acid sequence (Berthelsen *et al.*, 1996). On the basis of this peptide sequence, degenerate oligonucleotides were designed and used to amplify a specific PREP1 fragment from a human cDNA library. This was used as a probe to screen a HeLa cDNA library from which the human Prep1 cDNA was obtained (Berthelsen *et al.*, 1998a).

Like Meis proteins, Prep proteins can associate with Pbx in the absence of DNA, suggesting that Prep and Pbx, like Meis and Pbx, are stable partners *in vivo* (Berthelsen *et al.*, 1998b). As there are at least four *Pbx* genes, four *Meis* genes, and two *Prep* genes, the number of potential PBC complexes in vertebrates is quite significant.

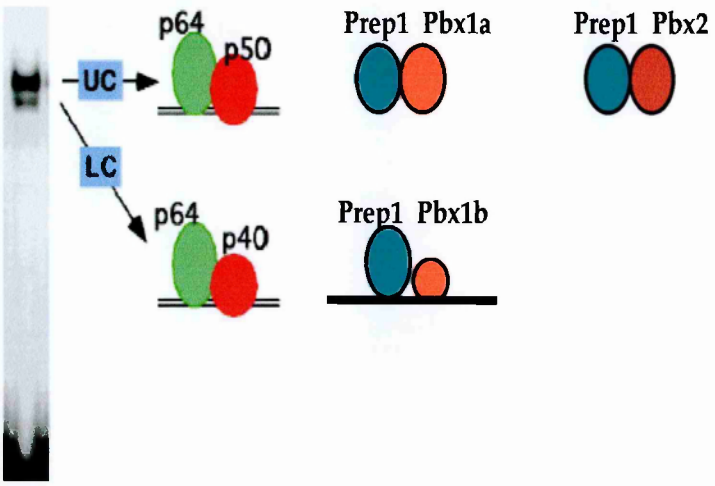


Figure 2.1
Subunit composition of UEF3.
The Urokinase Binding Factor (UEF3) binding the TGACAG sequence appears in EMSA as two retarded complexes: Upper Complex (UC) and Lower Complex(LC, which were later shown to be PREP1-PBX1A, PREP1-PBX1B and PREP1-PBX2 (Berthelsen *et at.*, 1996).

Figure 2.2

Panel A: sequence comparison of *r4- Hoxb1* enhancer between mouse and chicken. Capital letters emphasize identical bases and blocks with high similarity (block 1, repeat1, 2 and 3) are boxed.
Alignment of repeats 1-3 and comparison with Pbx consensus. The invariant bases are indicated in bold (from Popperl *et al.*, 1995).

Panel B: sequence of murine *Hoxb2* enhancer. Comparison between the r4- Hoxb2 and the Pbx/Hox consensus.

To directly address the role of Meinox-Pbx dimers as modulators of transcriptional activity of *Hox*, I analyzed the molecular interaction of Prep-Pbx dimers on two *Hox* targets: the *Hoxb1* and *Hoxb2* enhancers.

***Hoxb1 and Hoxb2* regulatory regions**

Developing mouse embryos (8.0 dp.c.) show an anterior expression limit of Hoxb1 in the hindbrain mapping at the boundary of rhombomere 3 and 4. Subsequently, at 9.5 d.p.c., Hoxb1 expression is up-regulated in the fourth rhombomere and down regulated in more posterior regions (Keyenes and Krumlauf, 1994). In later stages (11.5 d.p.c.) Hoxb1 expression is shut off in the hindbrain.

The spatially and temporally restricted expression of Hoxb1 in r4 is mediated by the Auto-Regulatory Element (*b1-ARE* element) through a positive auto regulatory feedback loop (Popperl *et al.*, 1995). The *b1-ARE* enhancer is able to drive the expression of a β-gal reporter gene in rhombomere 4 of transgenic mouse embryos. Likewise, in the developing *Drosophila* embryo, both Exd and the Hoxb1 homologue the Labial proteins are able to direct the expression of b1-ARE-LacZ in many subregions of the endogenous *labial* expression domain (Popperl *et al.*, 1995).

The *b1-ARE* enhancer is a *cis* regulatory element containing 148 bp and located in the 5' flanking region of the mouse *Hoxb1* gene. It contains three conserved sequence motifs, termed repeats 1, 2 and 3 (R1, R2, R3), which are similar to *Hox-Pbx* (PH) binding sites (Fig 2.2). These repeats are all necessary for *b1-ARE* activity, since mutations in a single motif have different affects; R2 makes the least contribution, while R3 seems to be the most crucial site for the activity. Three copies of this repeat cloned in front of the basal promoter are able to direct r4 restricted expression in mice. Thus, despite the sequence similarities among the sites, they are not functionally equivalent. Sequence comparisons in mouse, chicken and pufferfish revealed that the *b1-ARE* element is highly conserved in particular in the R1, R2 and R3 consensus sites and in block1 region (Fig 2.2). A functional conservation of this element is demonstrated by the ability to restrict the expression in pufferfish rhombomere 4 (Popperl *et al.*, 1995).

Group 2 paralogous genes (*Hoxa2* and *Hoxb2*) are both upregulated in r3 and r5 by the activity of the transcription factor Krox20. This is not surprising since the group originates from the duplication of an ancestral locus, resulting in a duplication of *cis*-regulative regions. However, in r4 the regulation of *Hoxa2* and *Hoxb2* is different, since while Hoxb2 is present, Hoxa2 is not (Krumlauf *et al.*, 1993), suggesting that different mechanisms are involved.

The specific expression of Hoxb2 in r4 is controlled by a *cis*-regulatory region containing a conserved *PBC/Hoxb1* binding site (Fig 2.2) (Maconochie *et al.*, 1997). This element is a sequence of 181 bp able to confer expression of a reporter gene specifically in r4 and associated neural crest, and capable of mediating a retinoid response, even if *RARE* response elements are not present (Maconochie *et al.*, 1997). This element contains a site (Hoxb2 r4 site, Fig 2.2) the sequence of which is very similar to the Pbx consensus, recognized by Pbx-Hoxb1 dimers (Fig 2.2). Mutation of this site results in he inability of the element to drive LacZ expression in rhombomere 4 (Maconochie *et al.*, 1997).

When taken together, the data suggest that the expression of Hoxb2 in r4 requires Hoxb1 and Pbx-Hoxb1 binding sites (Maconochie *et al.*, 1997).

CHAPTER 3
RESULTS AND DISCUSSIONI

Expression and structure of murine Prep1.

TALE homeodomain proteins include the PBC and Meinox subfamilies. Meinox protein members, including Hth, the four Meis genes products and the two Prep proteins Prep1 and Prep2, form heterodimer complexes with PBC proteins. Heterodimerization is crucial to DNA binding and nuclear translocation (Berthelsen *et al.*, 1999). PBC-proteins, such as Exd and Pbx factors, bind DNA in combination with HOX proteins, thereby modulating their DNA-binding specificity. For this reason, TALE proteins may play a crucial role in embryonic development and differentiation pathways.

The interaction of Meinox and PBC proteins is DNA-independent, requires specific conserved regions in the respective amino-termini and is compatible with the interaction of the PBC homeodomain with anterior Hox proteins.

In vivo, co-transfection of Prep1 cDNA, but not that of Prep2, increases Pbx-*Hoxb1* activated transcription from the b1-ARE sequence of the *Hoxb1* enhancer. (Berthelsen *et al.*, 1998; Fognani *et al.*, 2002)

Prep1 is always found in a DNA-binding complex with members of the Pbx protein family, the identity of which varies among different tissues and during embryogenesis. I have isolated murine Prep1 cDNA and analyzed expression of Prep1 both in adult and in embryonic murine tissues and observed that Prep1 is ubiquitously expressed in developing embryos as well as in adult mice. Furthermore, I found that Prep1, Pbx1, Pbx2 are present in the embryo extracts at the time when Pbx is acting on *Hox* regulation, which is compatible with a role for Prep1 as a regulator of Pbx-Hox interaction and function *in vivo*.

3.1 Sequence homologies of murine Prep1.

In order to study the role of the Prep1 *in vivo* I cloned the murine Prep1 gene. Using degenerate oligonucleotides and PolyA+mRNA isolated from murine fibroblast (NIH3T3 cells) I amplified a fragment of 300 bp by 3′ RACE (Fig 3.1). The 300 bp murine fragment showed 98% nucleotide sequence homology to the corresponding human PREP1 and was used to screen a 14.5 d.p.c. murine cDNA embryonic library (Ferretti *et al.*, 1999). Two overlapping cDNA clones of 2.4 and 3.4 kb) were isolated (Fig 3.1; Ferretti *et al.*, 1999). The 3.4 kb murine Prep1 cDNA contains 143 nucleotides of untranslated region (UTR), an open reading frame (ORF) of 1305 nucleotides encoding for 435 amino acids and 1971 bp of 3′ untranslated region (Fig 3.1). Besides the high sequence similarity within the coding sequences, both human and mouse cDNA show a long (2kb) 3′ UTR (Fig 3.1).

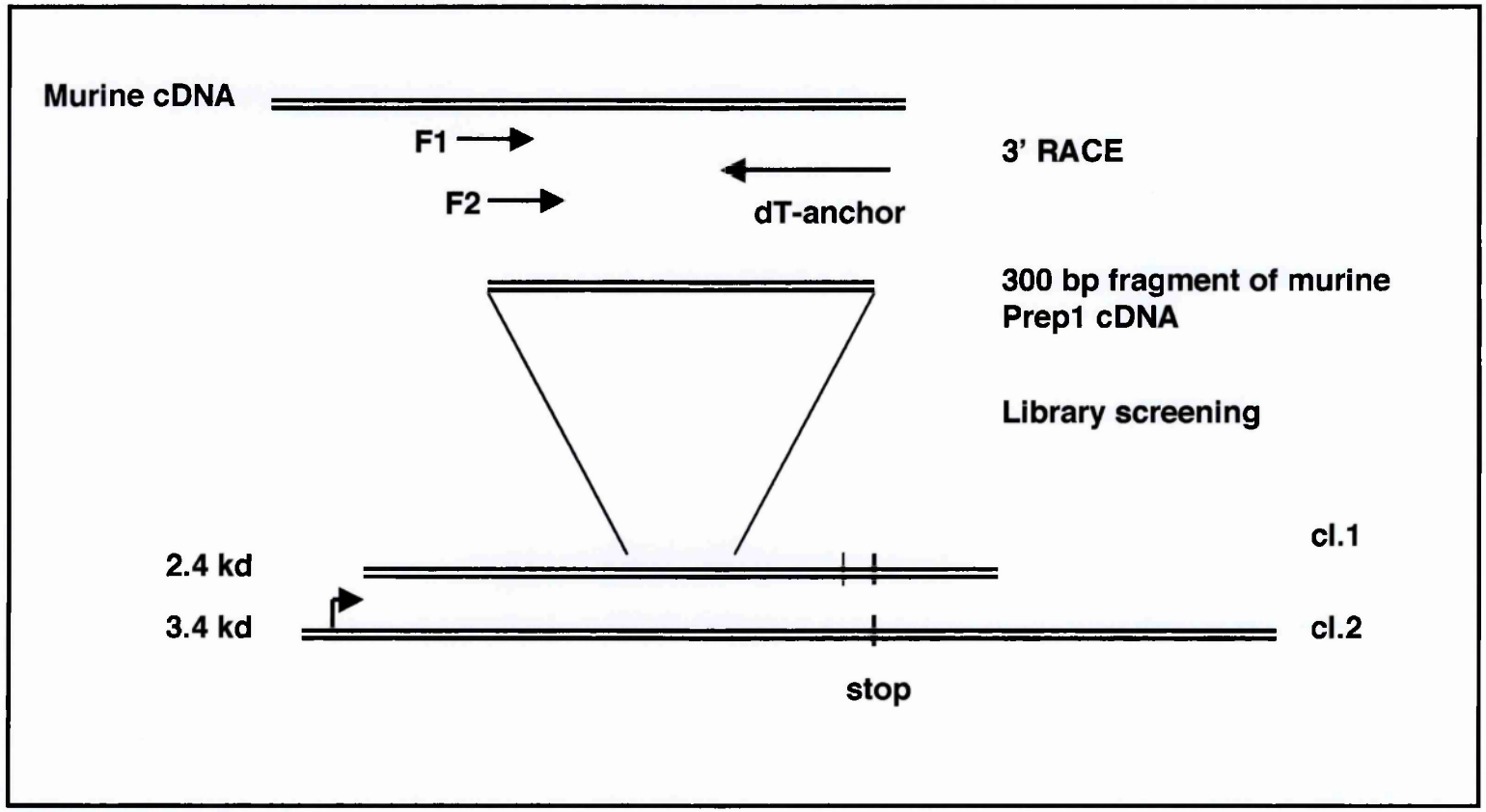


Figure 3. 1
Cloning of murine prep1 cDNA

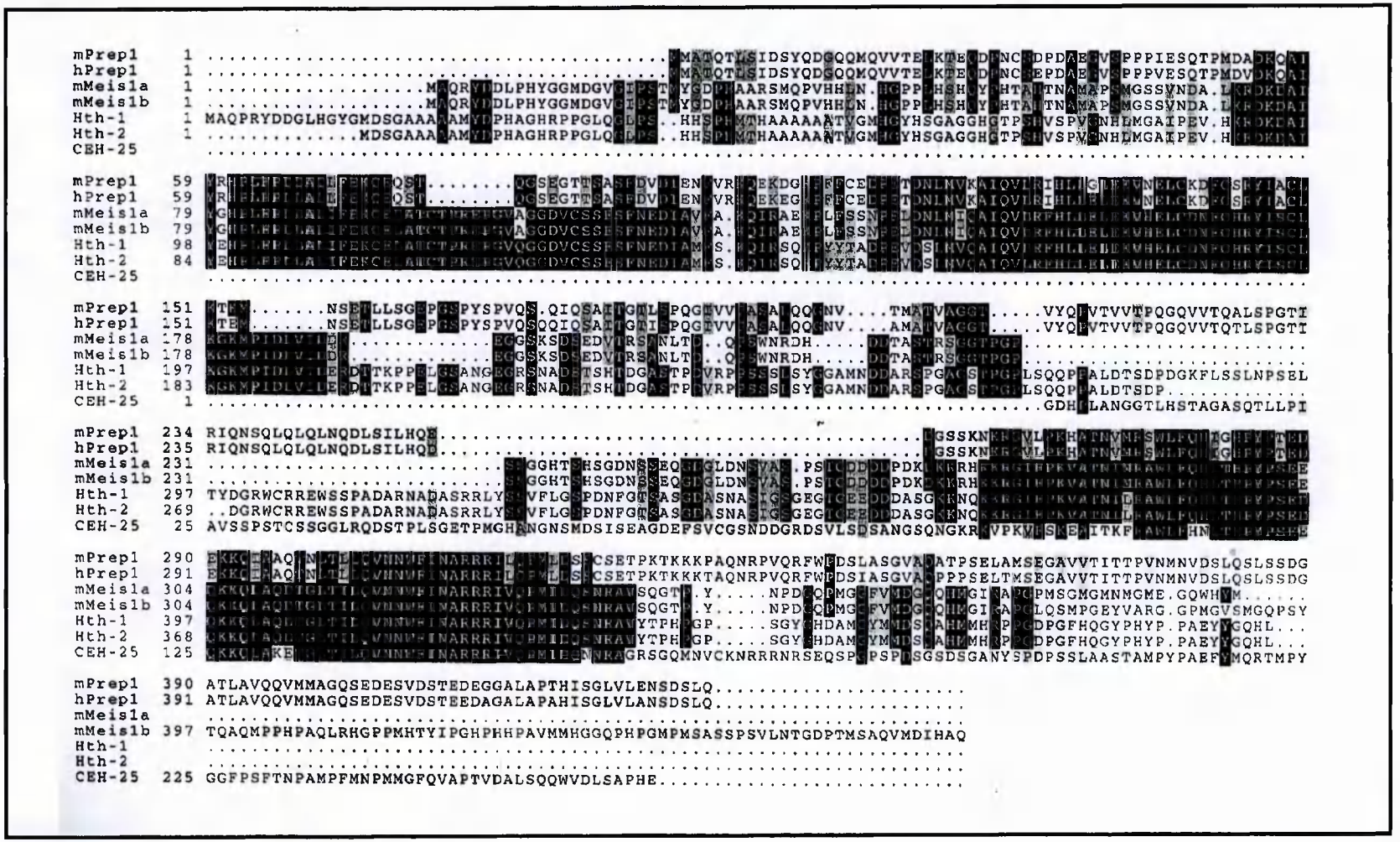


Figure 3.2
Comparison of the aminoacid sequences of murine/human Prep1 and Meinox proteins. The aminoacid identities are indicated in bold, The identity is particular high in the homeodomain HD (indicated with a blue double underline) and in the protein-protein interaction domains HR1 and HR2 (indicated with a single blue underline. Outside HR1, HR2 and HD there is no significant homology between Prep1, Meis and Hth. The dots represent gap in the sequences necessary to allow optimal alignment.

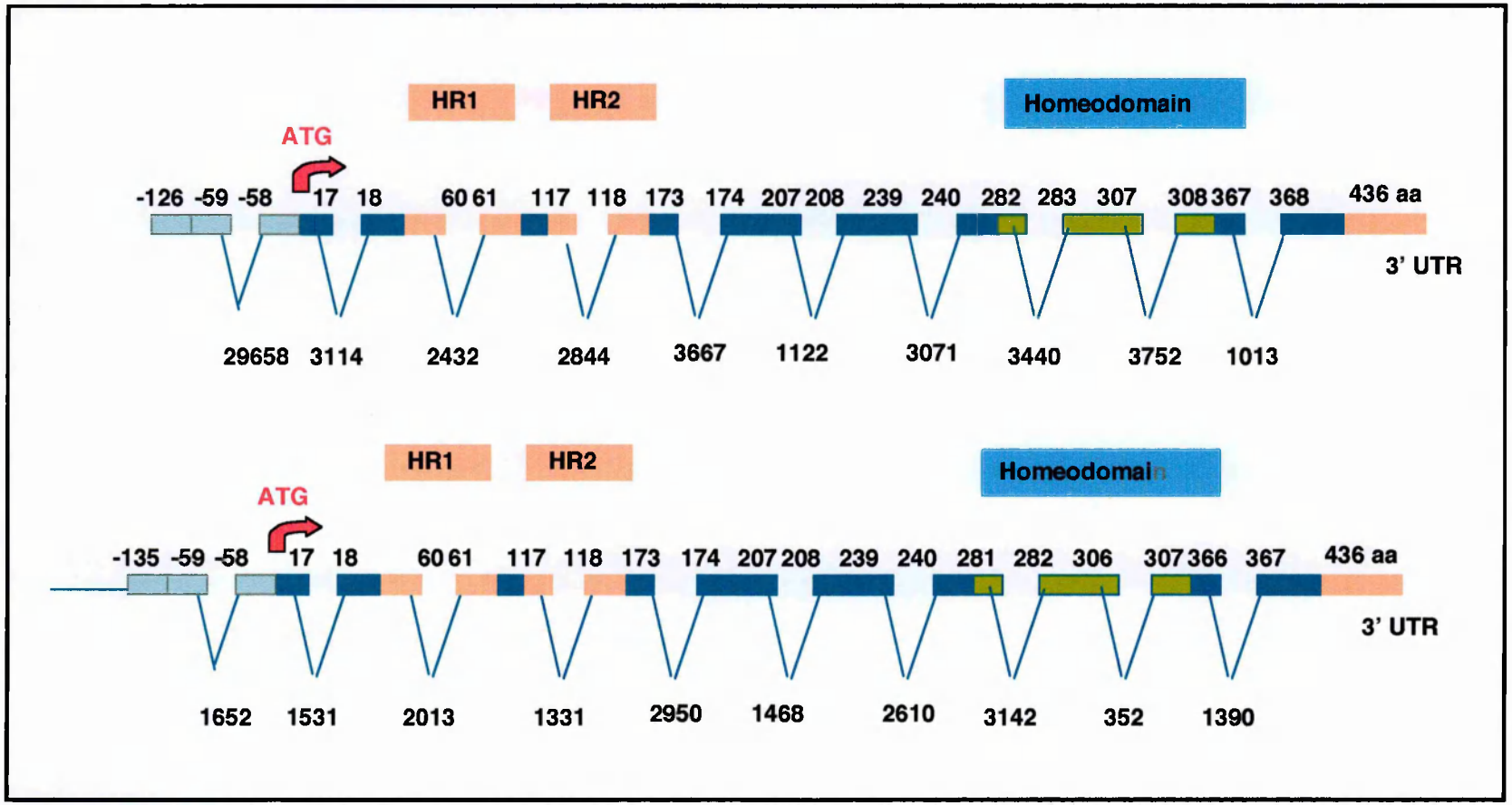


Figure 3. 3
Representation of exon-intron structure of murine Prep1 (lower diagram) and comparison with human PREP1 gene map (upper diagram).

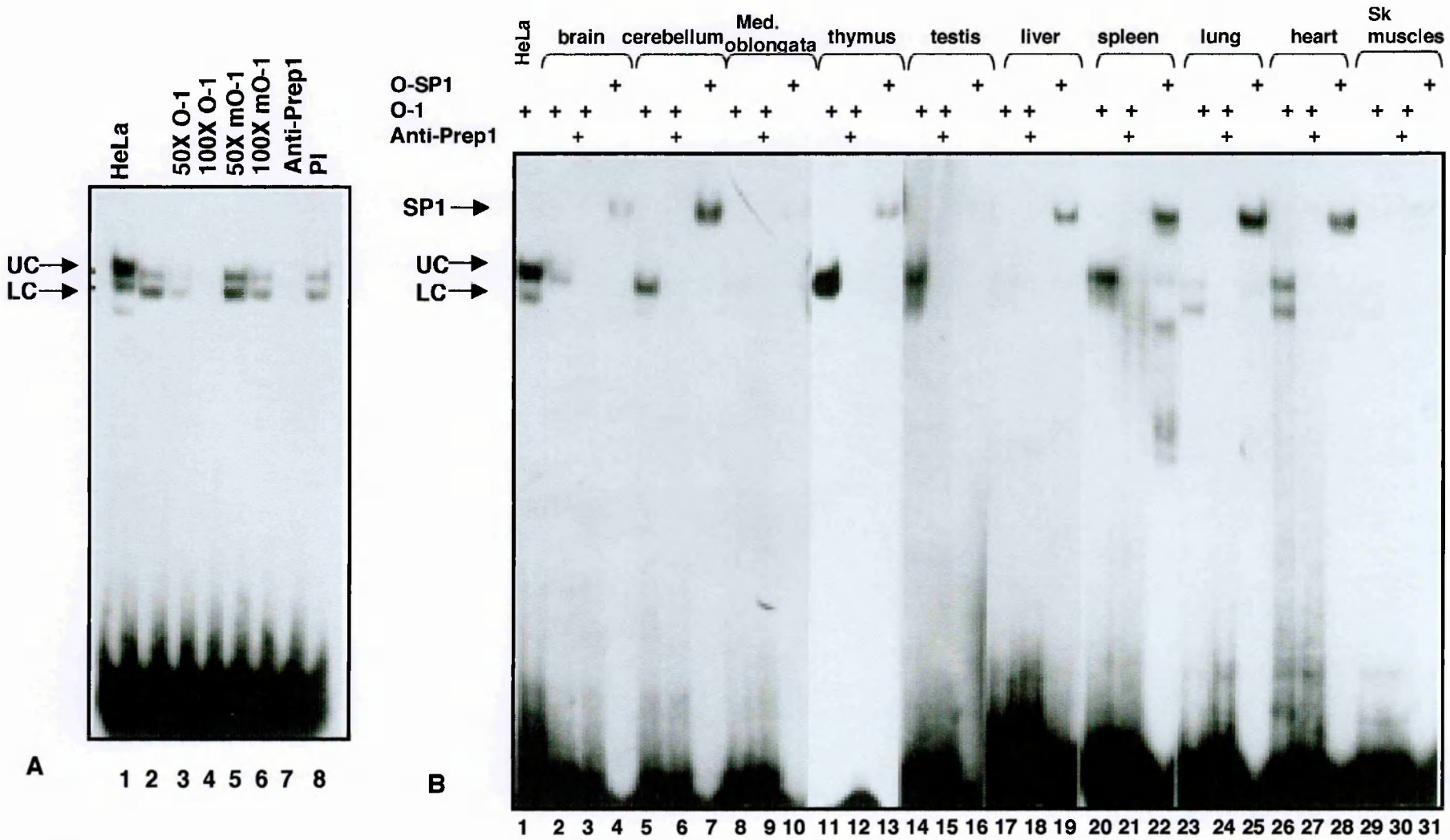


Figure 3.4

Prep1-Pbx heterodimers are present in murine extracts from various adult tissues.

A) EMSA analysis of NIH3T3 nuclear extracts using the O-1 oligonucleotide results in the formation of UC and LC. The binding specificity was analyzed using a competition assay where 50-or100 fold molar excess of n labeled wild type or mutated O1 was added.

B) EMSA analysis of nuclear extracts isolated from different adult mouse tissues using the O-1 oligonucleotide. O-SP1 is an oligonucleotide that binds the ubiquitous factor Sp1. Anti-Prep1 antibody was added to the binding reaction to test for the presence of Prep1 in the binding complexes.

Human and murine Prep1 share 100% identity in the homeodomain with 21 changes in 435 residues and one single codon deletion of the 20 changes, 10 result in conservative substitutions. Overall there is a similarity of 97% between the two proteins (Fig 3.2). Murine Prep1, like the human protein, contains a homeodomain belonging to the Meinox sub-family (Burling *et al.*, 1997), and two adjacent domains, located at the N-terminus and called Homologous Regions 1 and 2 (HR1 and HR2), which are required for the interaction of Prep1 with the PBC proteins (Berthelsen *et al.*, 1998b; Ferretti *et al.*, 1999).

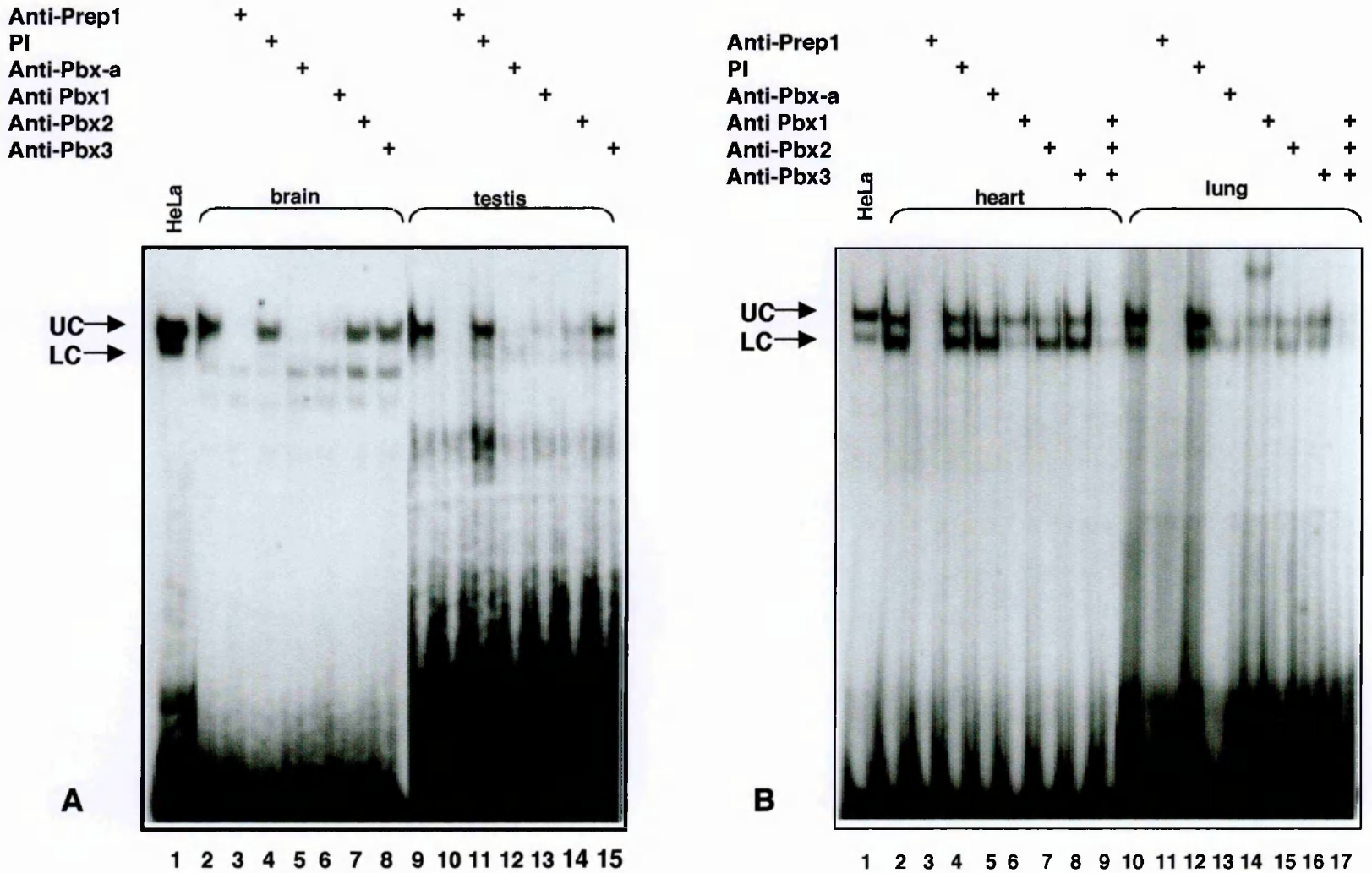
3.2 Human and murine PREP1 genes have an identical intron-exon organization.

Prep1 maps to human chromosome 21 in position q22.3 telomeric to Down Syndrome critical region and to murine chromosome 17 B/C (Berthelsen *et al.*, 1998d). I have searched the murine genomic data bank (NIH) for genomic sequences in order to identify the exon-intron structure of the murine Prep1 gene. I analyzed 11 unordered fragments and used this information to define the intron-exon organization and to determine the size of most of the murine Prep1 introns. The comparison between human and murine Prep1 genes reveled that the intron-exon organization was highly conserved (Fig 3.3).

3.3 Murine Prep1 is found in DNA-binding complexes with Pbx proteins.

Previous results obtained in our laboratory have shown that human nuclear extracts obtained from Hela cells contain DNA-binding Prep1-Pbx heterodimers (Berthelsen *et al.*, 1998b). Two retarded EMSA complexes, Upper (UC) and Lower Complexes (LC), have been characterized in HeLa nuclear extracts. They correspond to the PREP1-PBX2 and PREP1-PBX1B heterodimers. These complexes bind the TGACAG sequence contained in the O-1 oligonucleotide of the human *urokinase* enhancer, which was originally used by J. Berthelsen to purify the human Prep1 protein (Berthelsen *et al.*, 1996).

In order to test if the same kind of DNA-binding activities were also present in murine nuclear extracts, I performed EMSA on NIH3T3 nuclear extracts using the labeled O-1 oligonucleotide. As shown in figure 3.4 (panel A), two retarded bands are detected, the Upper and Lower Complex (UC and LC). These two retarded bands migrate similarly to the HeLa PREP1-PBX complexes (Fig 3.4 compare lane 1 and 2). Binding specificity was tested using an excess of unlabeled O-1: the binding activity was competed by unlabeled wild-type O-1 (Fig 3.4, lanes 3 and 4), but not a mutated form of O-1 (Fig 3.4, lanes 5 and 6). Moreover the formation of both complexes was strongly inhibited by incubation with anti-Prep1 antibodies (Fig 3.4, lane 7) while the pre-immune serum did not affect the binding activity (Fig 3.4, lane 8). To analyze the tissue-specific distribution of Prep1 in adult mice, I tested for the presence of the same DNA-binding activity in several adult mouse tissues. As shown in figure 3.4 (panel B), Prep1-Pbx heterodimers were found in all tissues tested (UC and LC). The binding activity was normalized using the o-Sp1 oligonucleotide specific for the ubiquitous nuclear factor Sp1 (Fig 3.4B, lanes 4, 7, 10, 13, 16, 19, 22, 25, 28 and 31). While Prep1-containing complexes were present ubiquitously, different ratios of the two retarded complexes, UC and LC are observed in different tissues.



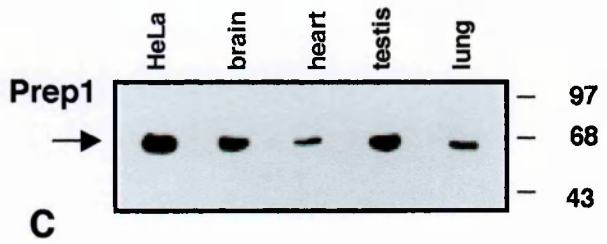


Figure 3.5

Analysis of subunit composition of the Prep1-Pbx complexes (UC and LC) in mouse tissues. EMSA was performed using the O-1 oligonucleotide and nuclear extracts in the absence or in the presence of anti-Prep1 and isoform specific anti-Pbx antibodies.

A) Brain and testis.

B) Heart and lung

C) Immunoblotting analysis of brain, heart, testis and lung murine nuclear extract using an anti-Prep1 antibody.

In some tissues, like thymus and spleen, only the UC is detected (Fig 3.4B, lanes 11 and 20). In other tissues (brain, cerebellum, medulla oblongata, and testis) both complexes were observed but with higher levels of UC (Fig 3.4B, lanes 2, 6, 8, 14). In muscle, heart and liver the two complexes are present in equal amounts (Fig 3.4 B lanes 17, 29 and 26). The binding specificity was demonstrated by incubating with anti-Prep antibodies, which abolished the formation of both heterodimers (Fig 3.4 B lanes 3, 6, 9, 12, 15, 18, 21, 24, 27 and 30). The antibodies are specific for Prep1 and do not recognize the homologous Meis proteins (Berthelsen *et al.*, 1998b).

3.4 The ratio of Prep-Pbx complexes reflects a differential expression of the Pbx components in various tissues.

To further investigate the nature of the Prep1 heterodimers in murine tissues, I analyzed the subunit composition of UC and LC in tissues having different ratios of the two complexes (Fig 3.5). To perform this analysis I used anti-Pbx antibodies that are able to distinguish between the different Pbx family members. I observed a tissue-dependent subunit composition of Prep-Pbx complexes, in which Pbx is the variable component.

For example, as shown in figure 3.5 A, nuclear extract from adult brain contains UC as the predominant form. Both anti-Prep1 and anti-Pbx-a antibodies block the binding of this complex to the O-1 oligonucleotide (Fig 3.5A, lanes 3, 5 and 6). The anti-Pbx-a antibody only recognizes the 50 kDa Pbx splicing isoform, suggesting that the heterodimers in the adult brain contain Prep1 and Pbx1a, Pbx2 or Pbx3a (Fig 3.5A, lane 5). Moreover, an anti-Pbx1, but not an anti-Pbx2 or anti-Pbx3 antibody, almost completely inhibits the formation of the UC band, suggesting that UC is composed of Prep1 and Pbx1a (Fig 3.5A, compare lanes 6, 7 and 8). This is in agreement with previously published data showing that Pbx1a is the predominant Pbx form expressed in rat brain (Monica *et al.*, 1996). In contrast, in testis I observed that the formation of the UC band is completely inhibited by an anti-Prep1 antibody but only partially by either an anti-Pbx1 or an anti-Pbx2 antibody. No effects are observed with an anti-Pbx3 antibody suggesting that Prep1-Pbx1a and Prep1-Pbx2 are present in the UC in mouse testis (Fig 3.5A). In addition, the weak LC is fully inhibited by anti-Prep1 and anti-Pbx1 antibodies, indicating that the LC in mouse testis is a Prep1-Pbx1b heterodimer (Fig 3.5A, lanes 3 and 6).

In heart extracts the formation of UC is partially inhibited by both anti-Pbx1 and anti-Pbx2 antibodies, suggesting that the complex contains Prep1 and Pbx1a or Pbx2 (Fig 3.5B, lanes 3, 6, 7). In the lung (Fig 3.5B lanes 11, 14, 15) both UC and LC bands are partly inhibited by anti-Pbx1, anti-Pbx2 and anti-Pbx3 antibodies and fully inhibited by the simultaneously addition of all three antibodies, showing that the lung UC is composed by Prep1-Pbx1a, Prep1-Pbx2 and Prep1-Pbx3a, while LC is formed by Prep1-Pbx1b and Prep1-Pbx3b (Fig 3.5B, lanes 16 and 17).

To verify that the formation of UC and LC did not depend on variation in the level of Prep1, I performed immunoblotting analysis on different tissue extracts using an anti-Prep1 antibody. In all cases I observed a 64 kDa band co-migrating with the human PREP1 band present in HeLa nuclear extract (Fig 3.5 C).

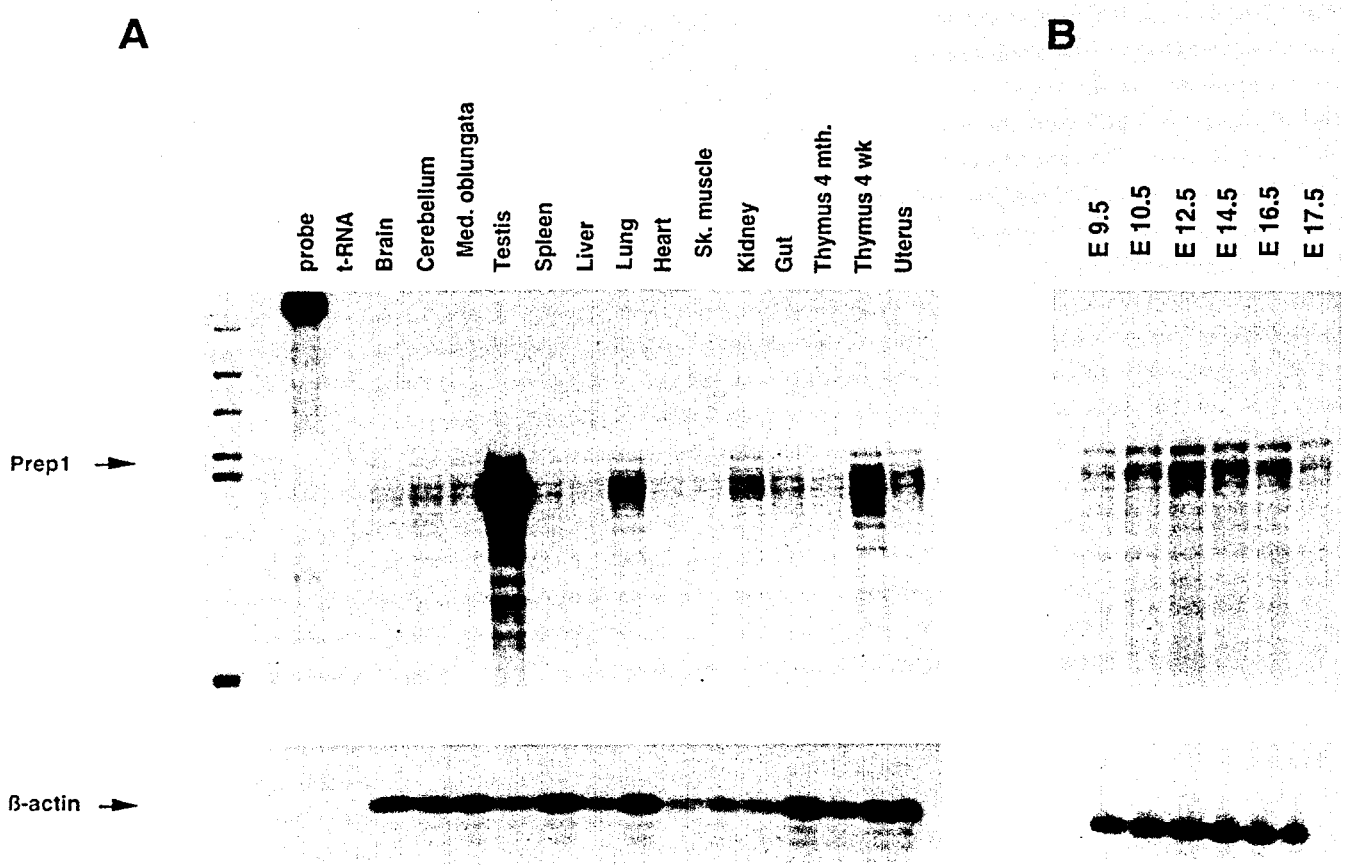


Figure 3.6

Prep1 expression in murine adult tissues and embryos.

A) RNAse protection experiments using mouse Prep1 oligo on total RNA purified from adult tissues

B) RNAse protection experiments using mouse Prep1 oligo to probe on total RNA purified from dissected embryos at (5-17.5 d.p.c. stages of development.

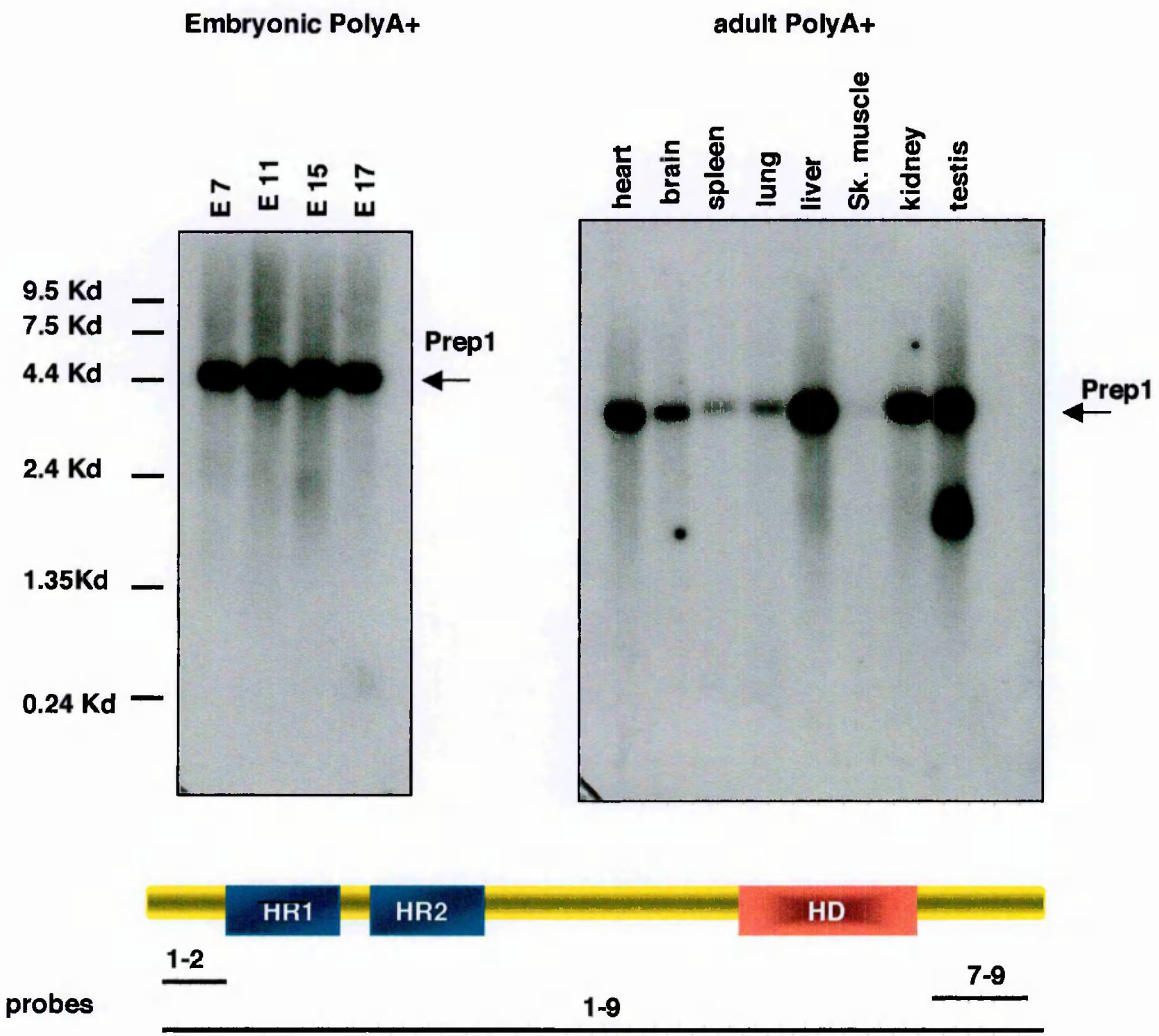


Figure 3.7
Northern Blot analysis of murine Prep1. Membranes containing 2μg/lane of PolyA+ RNA obtained from total murine embryos and adult tissues (Clontech, Palo Alto, CA) were hybridized with different Prep1 cDNA fragments as indicated in the panel.

In conclusion, the data show that Prep1-Pbx complexes are present in all murine adult tissues analyzed. However, we observe a tissue-type dependent subunit composition of the Prep1-Pbx complexes in which the Pbx isoforms differ between various tissues.

3.5 Prep1 expression in adult mice and mouse embryos.

Since Prep1 is expressed at a very low level, I used the more sensitive RNAse protection assay in order to analyze the expression pattern of murine Prep1. Total RNA from adult mice was assayed using a 210 bp probe located at the 3′ end of the homeobox domain in the murine Prep1 cDNA. The presence of a 210 bp protected fragment indicates expression of mPrep1 mRNA. A β-actin cDNA fragment was used as internal control. I found that Prep1 is expressed in all adult mouse tissues analyzed. Prep1 is very abundant in testis and thymus and lowest in skeletal muscle and heart (Fig 3.6A). In embryo Prep1 is already present at 9.5 d.p.c. and its expression remains constant from 9.5 d.p.c. to 17.5 d.p.c (Fig 3.6B). The multiple bands observed in Fig 3.6B.probably depend on the presence of a polyA stretch within the ribo-probe which is susceptible to RNAse cleavage. Alternatively, they may represent protection of different isoforms of Prep1 mRNA. However, Southern blots and mapping FISH suggest the existence of only one PREP1 gene isoform (Berthelsen *et al.*, 1998c). The multiple bands could also be due to the presence of other Prep family members.

In order to investigate the presence of mRNA splice variants I performed Northern blot analysis using PolyA$^+$ mRNA purified from adult and embryonic tissues.

Three different types of probe were used in these experiments (Fig 3.7):

1. the entire cDNA from initial ATG to stop codon.
2. a cDNA fragment located at the 5′ end, comprising 132 bp after the ATG.
3. a cDNA fragment located at the 3′ end, containing 240 bp before the stop codon.

As shown in Fig 3.7, I found one transcript of 4.0 kb in all tissues analyzed with the exception of testis where two transcripts of 4.0 kb and 2.0 kb were observed. This suggests the presence of a splice variant in testis or alternative polyadenilation site. No difference in the pattern was observed using any of the three probes. Northern blot analysis in adult tissues show high expression levels of Prep1 in liver, kidney and testis, low levels in skeletal muscle and intermediate levels in all other tissues analyzed (Fig 3.7).

Interestingly, the Prep1 mRNA expression pattern and levels reflect the level of Prep1-Pbx DNA binding activity, suggesting that Prep1 may be a limiting factor in Prep1-Pbx complex formation (compare Fig 3.7 with Fig 3.4).

3.6 Prep-Pbx complexes are present in mouse embryos.

In order to study the presence of Prep1 protein during embryonic development, I performed immunoblotting experiments using nuclear extracts obtained from dissected murine embryos at different stages of development. As shown in figure 3.8 A (upper panel), Prep1 protein is present in the embryo and its levels remain constant during development from 10.5 to 15.5 d.p.c.,

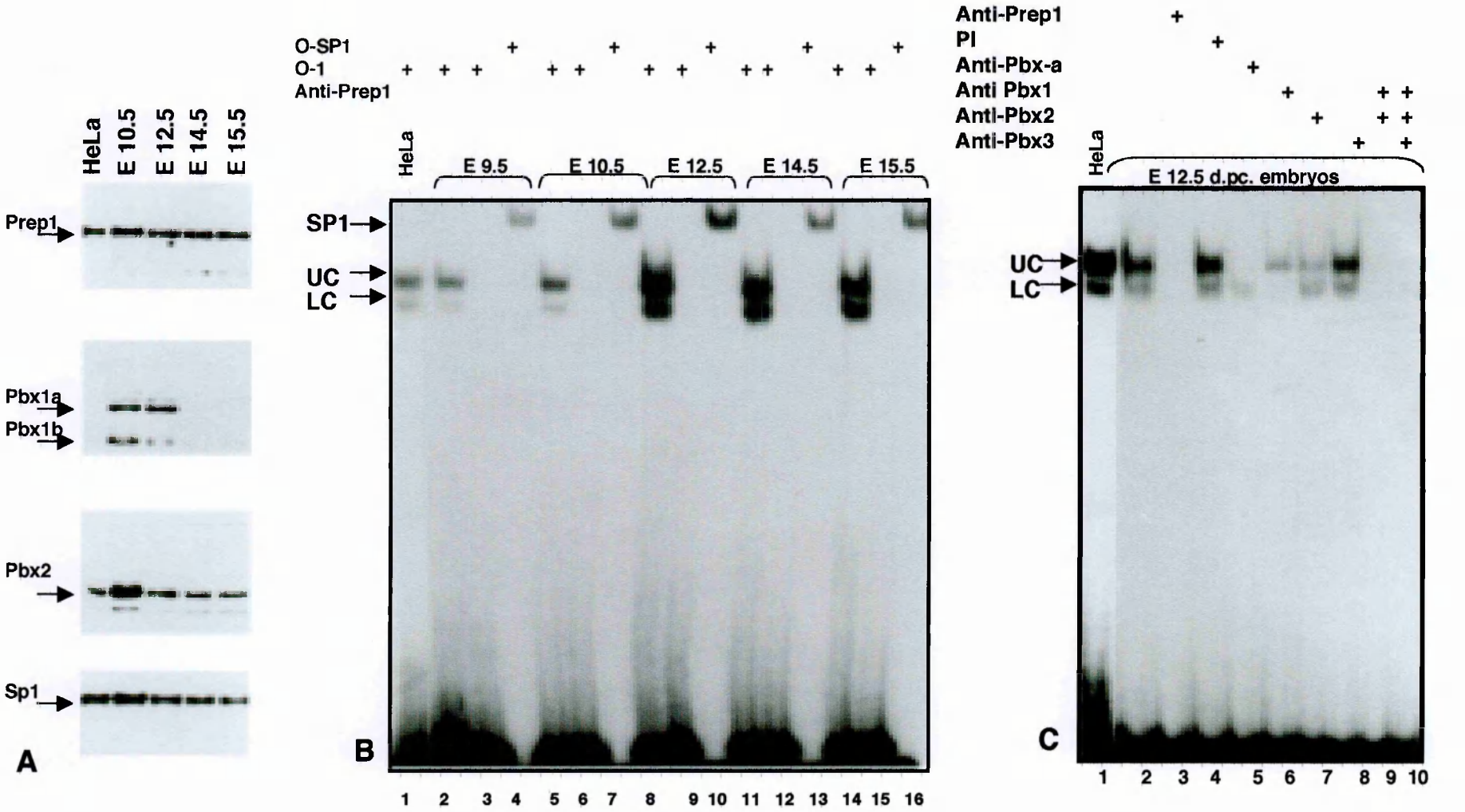


Figure 3.8

Prep1-Pbx complexes are present in whole mouse embryos.

A) Immunoblot experiment with nuclear extract purified from dissected embryos at 10.5-15.5 d.p.c. and tested with anti-Prep1, anti-Pbx-a, anti-Pbx2 and anti-Sp1 antibodies.

B) EMSA analysis using labeled O-1 oligonucleotide as a probe and embryonic nuclear extract obtained from 9.5-15-5 d.p.c. mice. As a control I used the O-Sp1 oligonucleotide containing the binding site for the ubiquitous factor Sp1.

C) The UC and LC composition at 12.5 d.p.c. in nuclear extract obtained from entire murine embryos was analyzed by EMSA using anti-Prep1, anti-Pbx1, anti-Pbx2 and anti-Pbx3 antibodies, as indicated.

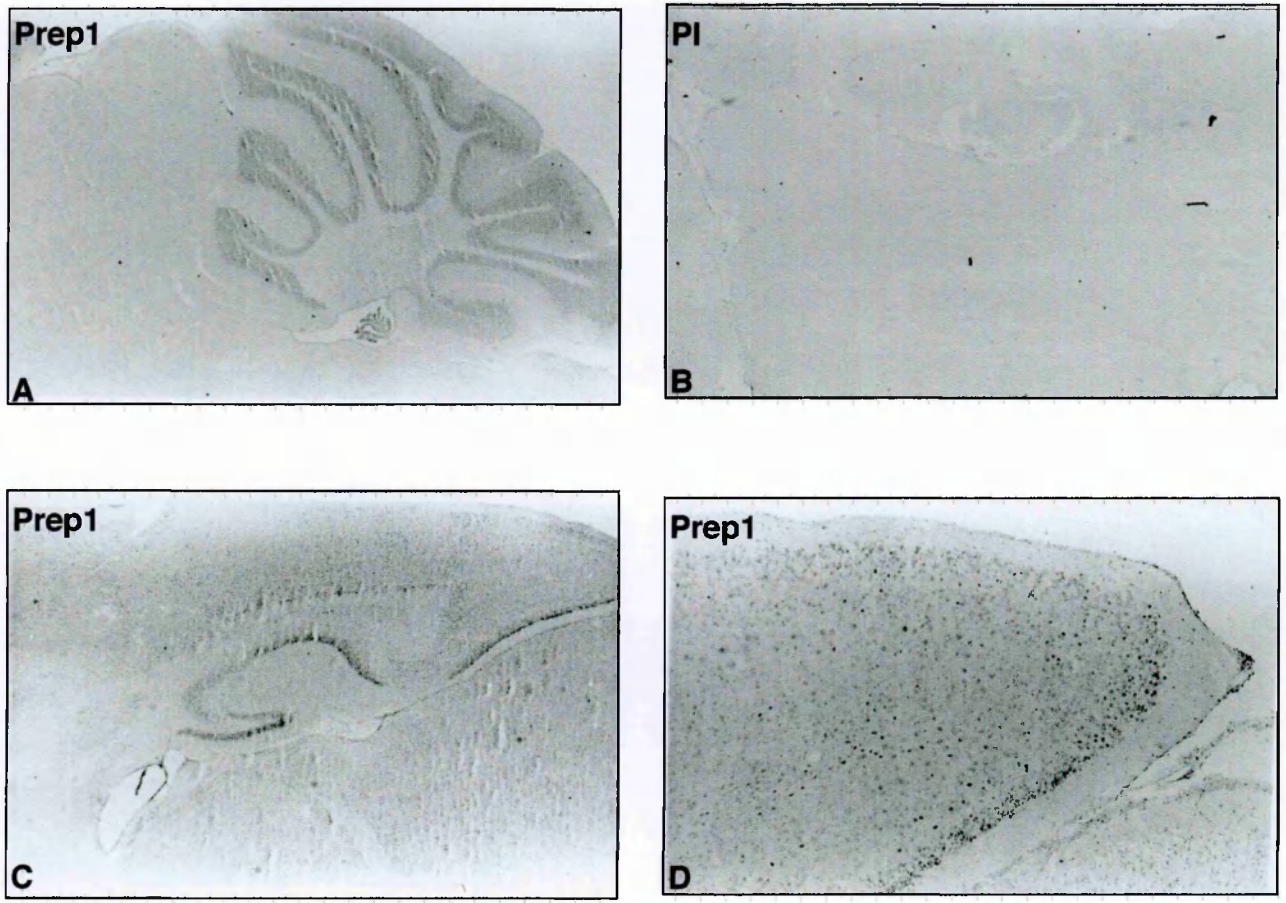


Figure 3.9
Immunohistochemical localization of Prep1 in murine adult brain (4 weeks) using Prep1 antibody (A, B and D) or pre-immune serum PI (B).

A) Localization of murine Prep1 in the cerebellum
B) Pre-immune serum
C) Localization of murine Prep1 in the hippocampus
D) Localization of murine Prep1 in the inferior colliculus

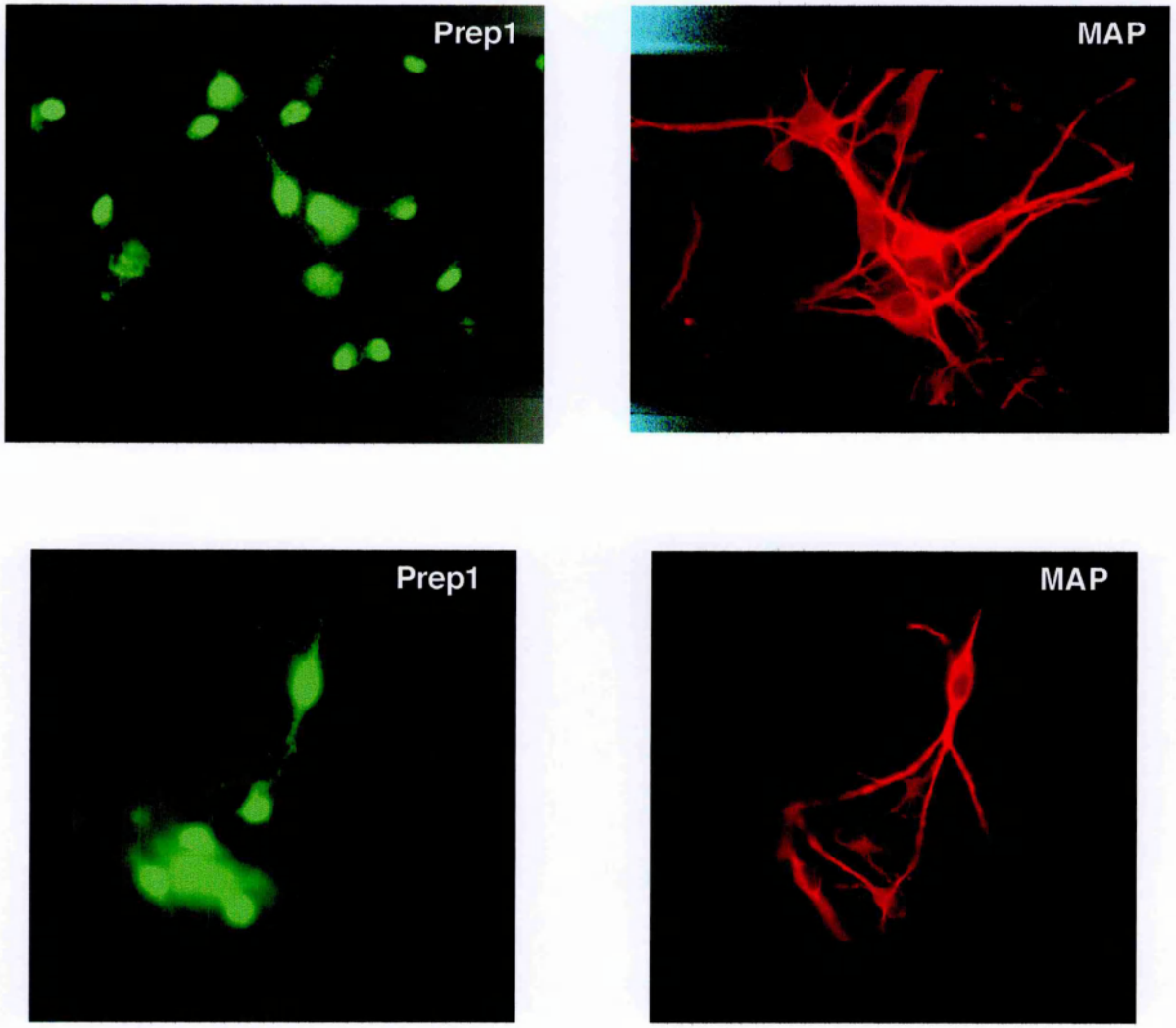


Figure 3. 10
Immunohistochemical localization of murine Prep1 in rat hippocampus neurons using Prep1 antibody (A, C) or anti-MAP antibodies -(C, D). Prep1 staining is revealed by fluorescin, while the control MAP protein is visualized by rhodamine staining.

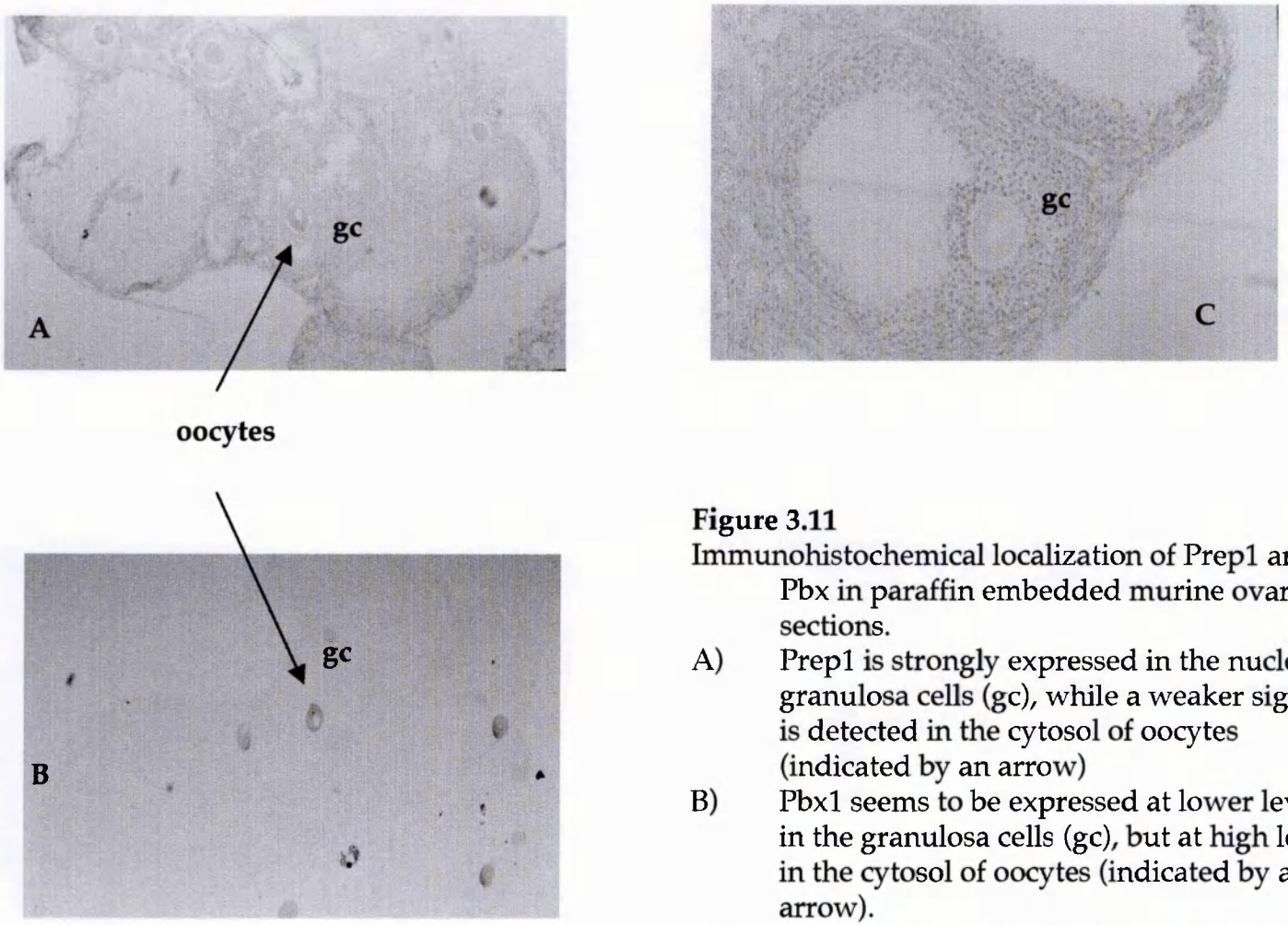


Figure 3.11
Immunohistochemical localization of Prep1 and Pbx in paraffin embedded murine ovary sections.

A) Prep1 is strongly expressed in the nuclei of granulosa cells (gc), while a weaker signal is detected in the cytosol of oocytes (indicated by an arrow)

B) Pbx1 seems to be expressed at lower level in the granulosa cells (gc), but at high level in the cytosol of oocytes (indicated by an arrow).

C) High magnification of late stage folliculus showing high expression of Prep1 in the granulosa cells (gc) and the presence of prep1in the cytoplasm of oocytes. No immunoreactivity was observed using a pre-immune serum (PI) (data not show).

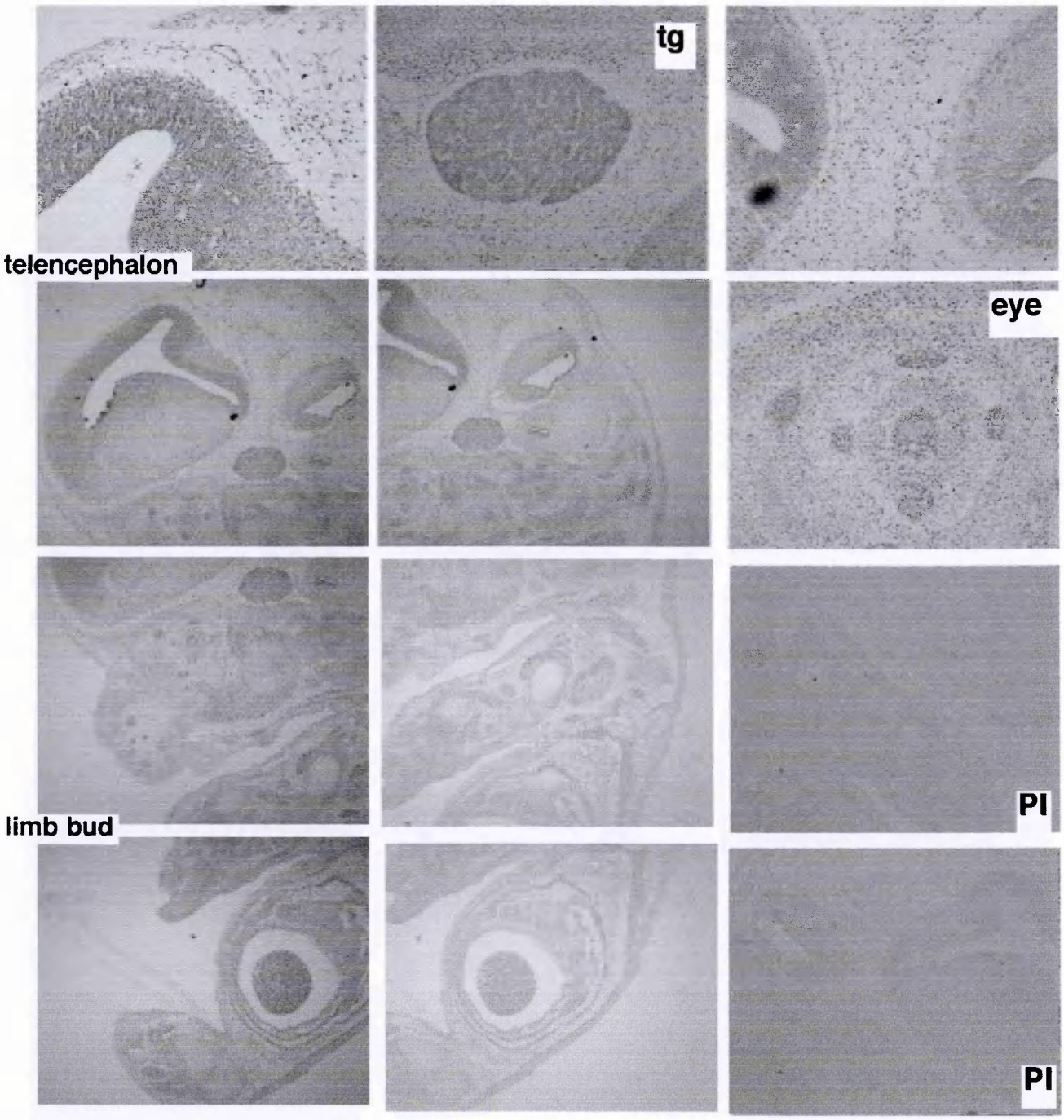


Figure 3.12

Prep1 expression during mouse embryogenesis

Sagittal sections of 13.5 d.p.c. mouse embryos were analyzed by immunohistochemical analysis with anti-Prep1 antibody. Widespread expression of Prep1 is seen in the nuclei. Areas of intense signal include telencephalon, and hind brain, tigeminal ganglion (tg) and eyes.

Note the absence of signal in control experiments performed using pre-immune serum (PI)

confirming the RNA expression data. I extended the analysis to various forms of the Prep dimerization partners using antibodies against Pbx proteins.

Figure 3.8 A shows that Pbx1a and 1b are clearly present at 10.5 and 12.5 d.p.c, but decline at later stages below the detection limit. Pbx2, on the contrary, is constitutively expressed at all stages. An antibody against the ubiquitous transcription factor Sp1 was used as a control.

Subsequently I looked for the presence of DNA-binding Prep-Pbx complexes (UC and LC) at different stages of embryonic development. Figure 3.8 B (lanes 2, 5, 8, 11 and 14) shows that Prep-Pbx complexes are formed at all stages tested. The binding intensity increases from day 12.5 d.p.c. when compared with Sp1 binding.

In order to identify Prep1 embryonic partners in UC and LC, I used anti-Prep and anti-Pbx-a antibodies for competition assays (Fig 3.8 C). Using anti-Prep1 and anti-Pbx antibodies I demonstrated that while anti-Prep1 prevents the binding of both complexes, anti-Pbx1 inhibits all LC formation (Fig 3.8 C, lane 6)., but only half of UC formation. Anti-Pbx2 antibody inhibits half of UC (Fig 3.8 C, lane 7). However, together anti-Pbx1 and anti-Pbx2 completely inhibit UC formation (Fig 3.8 C, lanes 9 and 10). These data suggest that UC is composed by both Prep1-Pbx1a and Prep1-Pbx2 heterodimers, while the UC only contains Prep1-Pbx1b heterodimers.

3.7 Prep1 is widely expressed in the nuclei of neuron of murine adult brain.

Prep1 protein is present in the nuclear extract from murine adult mouse brain. In order to test the cell-specific distribution of Prep1 in the adult brain, I performed immunohistochemical staining using brain sections from 4 week old mice (Fig 3.9). Prep1 appears to be present in the nuclei of cerebellum, in particular in the Purkinje cells (Fig 3.9A) and is also expressed in the hippocampus (Fig 3.9 C). Moreover, there are several nuclei expressing Prep1 in the inferior colliculus (Fig 3.9, panel D). Prep1 seem to have predominant staining in neuronal cells. Furthermore immunohistochemical localization of Prep1 in primary cultured hippocampal rat cells shows a strong nuclear staining. Figure 3.10 shows that Prep1 is present in the cellular body but not in axons. On the contrary, the microtubulin associated protein (MAP) used as a control specifically stained the cytoplasm of axons.

3.8 Prep1 shows cytoplasmic expression in oocytes of adult mice.

Immunohistochemical studies performed on paraffin embedded sections of different murine tissues show a predominantly nuclear staining for Prep1. However I did find one tissue in which Prep1 was localized in the cytoplasm: the oocytes. As shown in figure 3.11 Prep1 is expressed at high levels in the nuclei of granulosa cells (cg) surrounding the oocytes. On the contrary Pbx1 show a predominant staining in the cytoplasm of oocytes and weaker staining in the granulosa cells. However, additional experiments are required to further characterize the role played by the interaction of Prep-Pbx during ovary development.

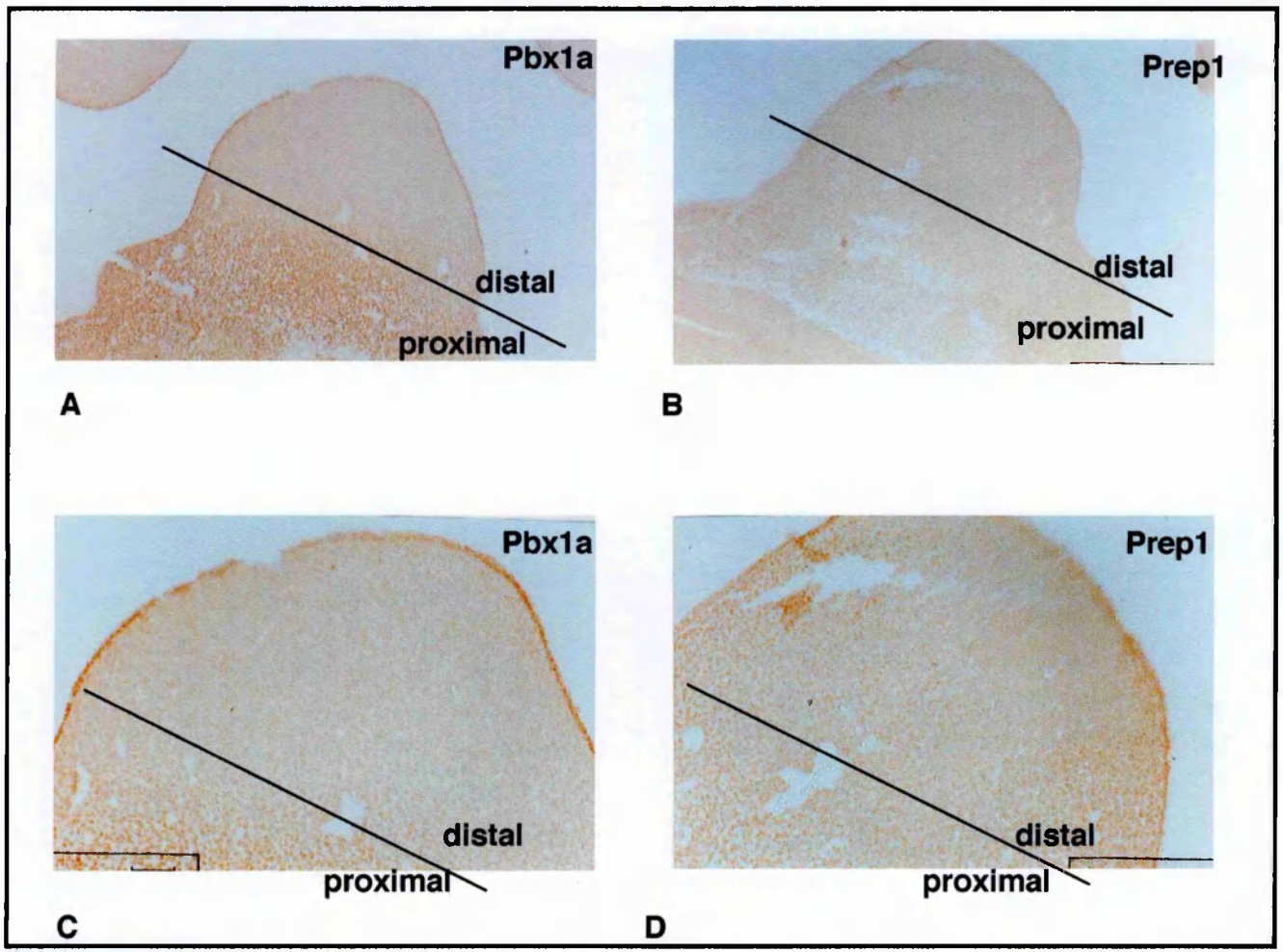


Figure 3.13 Prep1 and Pbx1 localization in limb bud of 10.5 d.p.c. mouse embryos.
(A, C) Immunohistochemical experiments using Pbx1 antibody show cytosolic expression in the distal part of the limb bud. On the contrary the staining is nuclear in the proximal part of the limb bud.
(B, D) Immunostaining with anti Prep1 antibody shows nuclear the staining in the limb bud. Prep1 does not show difference in term of nuclear/cytoplasmatic cellular localization along the distal axis.

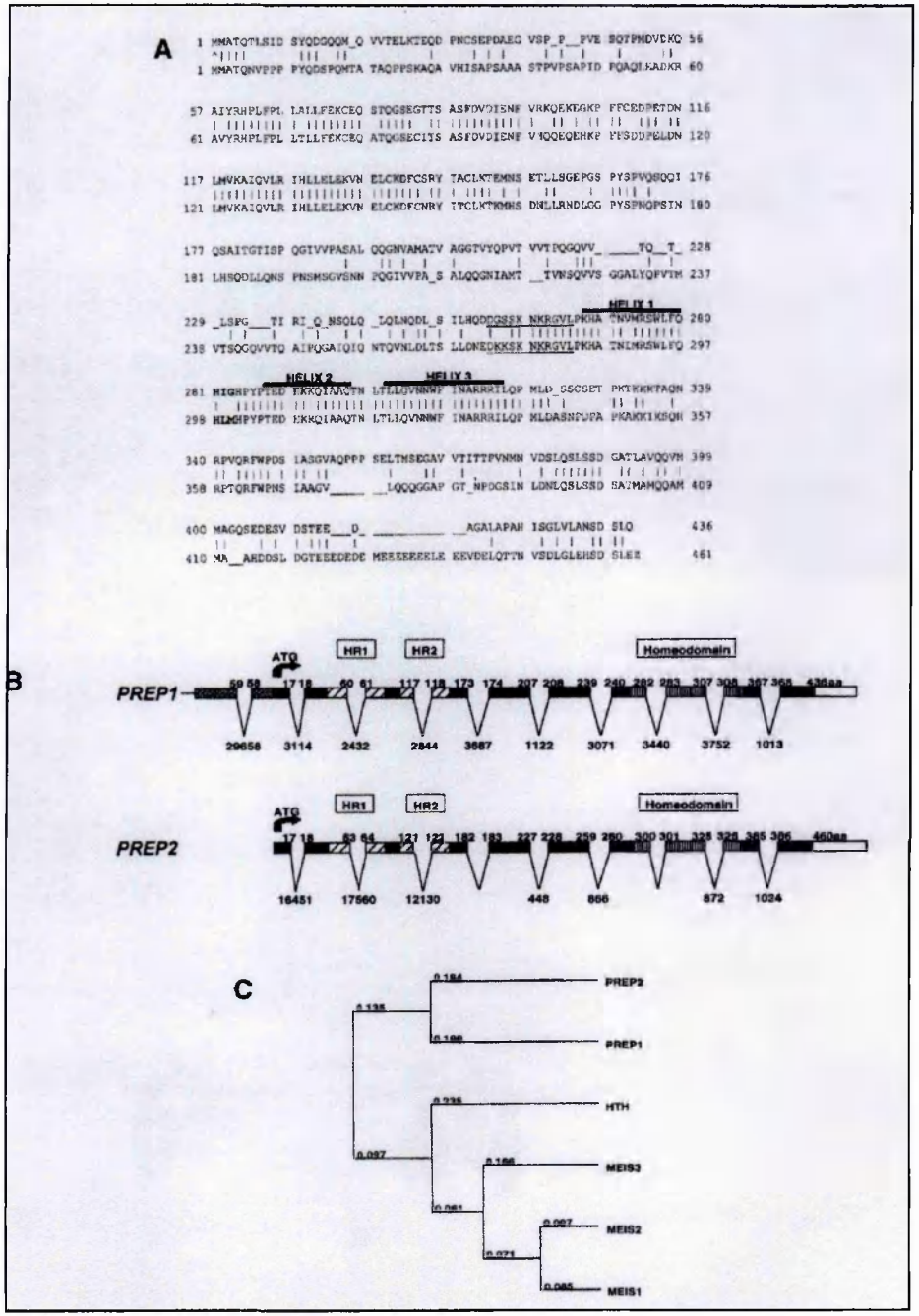


Figure 3. 14
Schematic representation of protein sequence and exon-intron organization of humanPREP1 and PREP2.
Panel :alignment of human PREP1 with human PREP2 proteins shows 52 % identity with human PREP1. Conserved residues are indicated with double underline, and single denotes the protein protein interaction region (HR1 and HR2) conserved among the Meinox family.
Panel B: Structural organization of human PREP1 and PREP. PREP1 maps on chromosome 21, while a 200 Kb fragment containing PREP2 maps on chromosome 11. The size of some PREP2 introns is not indicated because the sequence of some sub-fragment lacked overlapping verification. The number above refer to the amino acid residues in the protein. The number at the bottom indicate the intron size.
Panel C: phylpgenetic tree derived from comparison of protein sequences of Drosophila Hth, human MEIS1, MEIS2, MEIS3, and human PREP1 PREP2 . Clustalx and Nijplot applications were used. Number indicated branch lengths.

3.9 Prep1 is widely expressed during the embryonic mouse development.

The presence of Prep1 protein in 10.5 to 16.5 d.p.c. murine embryos was analyzed by immunohistochemical staining of paraffin embedded sections. As shown in Fig 3.12 and 4.1 Prep1 is broadly expressed during murine embryonic development and localizes to the nuclear compartment of the majority of the cells. At 10.5 d.p.c. Prep1 is present in the nuclei of developing brain, somites and limb buds (Fig 3.12 and 4.1).

At 13.5 d.p.c., Prep1 continues to be expressed in the brain and olfactory bulb where it is coexpressed with Pbx1. Prep1 is still present in the somites, dermomyotome and mesenchyme of the intervetebral discs. A positive signal is also detected in the chondrocytes, correlating with Pbx1 expression and function in the endochondral skeleton (Schnabel *et al.*, 2001; Selleri *et al.*, 2001). Prep1 is also present in the dorsal root ganglia and in neural crest derived structures (Fig 3.12). In conclusion, in mouse there is a strong correlation observed between Prep1 gene expression and nuclear localization of Pbx1.

3.10 Identification and structural characterization of a novel PREP1 paralogous gene: PREP2.

Recently, by homology searches of human genomic databases, we and others have identified and cloned a novel Meinox family member displaying a high degree of similarity with Prep1, which was termed PREP2 (Imoto *et al.*, 2001; Fognani *et al.*, 2002).

PREP2 is located on chromosome 11q24 and is expressed in a variety of human adult tissues but it displays a more restricted expression pattern than PREP1. PREP2 is longer than PREP1 (461 versus 435 aminoacids). The amino acidic residues in the two proteins are 52% identical with conservative substitutions that raise the overall similarity to 60%. The homeodomain and the two protein-protein interaction domains (HR1 and HR2) share the highest level of similarity while the C-termini is largely divergent (Fig 3.14 A and Fognani *et al.*, 2002).

In order to identify the exon-intron structure of the human PREP2 gene I have searched the human genomic data bank (NIH) for genomic sequences (Fig 3.14 B). Alignment of PREP2 cDNA with genomic data bank resulted in the identification of a 200 Kb sequence located on chromosome 11 that could be used to define the intron-exon organization of the PREP2 gene. Figure 3.14 B shows that human PREP1 and PREP2 have the same intron-exon organization, suggesting that intron-exon organization is highly conserved in the Prep family.

The generation of a phylogenetic tree using the CLUSTALX and NJPLOT software confirmed that PREP1 and PREP2 define a sub-family of Meinox proteins distinct from Meis and Hth (Fig 3.14 C and Fognani *et al.*, 2002).

The structural analysis of Prep2 reveals similar functional similarity with Prep1, however the presence of a poly-glutamate tail in Prep2 only, hints at functional diversification.

Discussion chapter 3

Human Prep1 is a homeodomain containing protein of the Meinox family capable forming heterodimers with the PBC protein family (Berthelsen *et al.*, 1998).

The homeotic transcription factors *Hox* have the unique property of directing the formation of the specific embryonic structures. The mechanisms by which field-specific selector proteins direct the development of these structures are not well understood. Furthermore, most *Hox*-target genes are unknown.

The fact that the Hox proteins bind DNA consensus sequences with low affinity and specificity adds complexity to the system. However, it is well known that cofactors such as Exd and PBC proteins can drive the Hox proteins to specific target sequences by increasing their DNA-binding affinity as well as selectivity. The PBC proteins therefore provided the initial key to understanding how the homeotic proteins work.

The identification of Meinox proteins that can regulate the PBC proteins introduces another level of complexity in Hox regulation and function. In fact, the Meinox proteins are able to influence not only the PBC proteins but also the Hox-PBC complexes. Transfection experiments have shown that Prep1 is able to activate the Pbx-Hoxb1 dependent *Hoxb1* promoter *b1-ARE* (Berthelsen *et al.*, 1998).

Prep1 gene is a member of the Meinox family

I have isolated the murine Prep1 gene, a member of the Meinox family, and analyzed its pattern of expression in adult mice and mouse embryos.

The cDNA and amino acid sequences show that murine Prep1 is highly homologous to its human orthologue. Murine Prep1 shares 100% identity with human PREP1 in the homeodomain and 95% similarity throughout the whole protein (Fig 3.2). Prep1 shares a high level of homology with other members of the Meinox family essentially only in the homeodomain and in two other regions. These regions are known as the homologous regions 1 and 2 (HR1 and HR2) and are located at the N-terminus. The non-homologous regions may be responsible for the functional differences observed between Prep and other Meinox proteins (Fig 3.2).

PREP1 maps to human chromosome 21 in position q22.3 telomeric to Down Syndrome critical region and to murine chromosome 17 B/C (Berthelsen *et al.*, 1998d). Moreover, the exon-intron organization of the two genes is highly conserved, suggesting that a positive selective pressure has maintained these genes during evolution (Fig 3.3).

Prep1 is widely expressed in adult tissues.

Both northern and immunoblotting experiments show that Prep1 is widely expressed in adult tissues with the highest levels seen in testis and thymus (Fig 3.6 and Fig 3.7). However, immunohistochemical experiments show a cell specific expression. For example, in mouse adult brain, Prep1 is expressed

at higher levels in cerebellum than in cerebral cortex and the staining seems to be specific for neuronal and not for glial cells (Fig 3.9). Furthermore, Prep1 is highly expressed in the hippocampus of adult mice, showing a strong nuclear signal in isolated hippocampal neurons (Fig 3.10).

Prep1-Pbx complexes show a tissue-type dependent subunit composition in different murine tissues.

EMSA experiments show that in these tissues Prep1 is the predominant partner of Pbx proteins, resulting in the formation of Prep1-Pbx complexes in the nucleus. While Prep1 seems widely present, the tissue-type dependent choice of Pbx partner could confer diverse activities to the complex. Prep1 can bind both the alternatively spliced forms of Pbx. I have shown that only some tissues contain Prep1-Pbx-b (LC), suggesting differential regulatory properties for UC and LC complexes (Fig. 3.4). The products of the two splice variants of Pbx1, Pbx1a and Pbx1b, show differential activity in combination with Hox proteins (Di Rocco *et al.*, 1997). However, the activity of the various Prep-Pbx complexes remains to be analyzed.

I have also shown that in the murine adult brain Prep1-Pbx1 is the predominant heterodimer. The absence of Hox protein in the rostral brain suggests a role for Prep1/Pbx1 in the regulation of brain activity in a Hox independent manner. In adult murine testis the UC is composed by Prep and Pbx-a splice variant (Fig. 3.5). The partial inhibition of this complex by the addition of both anti-Pbx1 and anti Pbx2 antibodies may suggest that this dimer also contains Pbx4. In heart the two dimers present correspond to Pbx1-Prep and Pbx2-Prep. Finally, in the lung, three combinations of dimer are present namely Prep1-Pbx1, Prep1-Pbx2, Prep-Pbx3.

Prep1 is expressed in the nuclei of adult murine tissues.

Prep1, and indeed all the members of Meinox protein family, seem to constitute a novel group of homeodomain proteins. They are able to form heterodimers and recognize the atypical DNA sequence TGACAG, thus suggesting that Prep1 and other members of the Meinox family may have overlapping functions. When present together the Meinox proteins are probably competing for the Pbx partners or additional regulative mechanisms could selectively modulate the affinity for Meinox versus PBC proteins. This adds another level of complexity, to Pbx-Meinox functional regulation.

One of the known mechanisms of regulation is nuclear/cytoplasmic localization.

Meinox proteins control the activity of PBC proteins by promoting their nuclear transport, both in cell culture and *in vivo*, which is absolutely required for their function (Mann and Abu-Shar, 1996; Rieckhof *et al.*, 1997; Pai *et al.*, 1998; Berthelsen *et al.*, 1998). In mouse, Pbx1 is present in the nuclei of proximal limb bud cells while it resides in the cytoplasm of distal limb bud cells (Fig 3.13). This pattern is consistent with the proximal pattern of expression of Meis1-2 proteins. The uniform expression of Prep1 in limb bud could suggest that it is not directly involved in controlling nuclear/cytoplasm translocation in the developing limb

bud, suggesting that the Meis proteins play an essential role in vertebrate limb development (Fig 3.13).

There is considerable evidence supporting the idea that PBC proteins must be localized in the nucleus in order to be active. The presence of both Prep1 and Pbx proteins in the cytoplasm of oocytes may depend on the inactive transcriptional status of these cells (Fig 3.11). We can hypothesize that after the fertilization, when the oocytes actively transcribed gene, Prep1-Pbx will translocate into the nucleus.

Prep1 is expressed in the embryo at different stages of development.

Using RNAse protection, northern and immunoblot analysis I have shown that Prep1 is expressed in the embryo at different stages without changes in pattern in its expression (Fig. 3.6 and 3.7).

Prep1 is broadly expressed during murine embryonic development and localizes to the nuclear compartment of the majority of the cells showing strong correlation observed between gene expression and nuclear localization of Pbx proteins. In the embryo Prep1 is present in the nuclei of developing brain, olfactory bulb, somites, dermomyotome and mesenchyme of the intervetebral discs, limb bud, dorsal root ganglia and in neural crest derived structures (Fig 3.12).

Could Prep1 be the limiting factor in heterodimer formation?.

Using immunohistochemistry I have also shown that while the level of Prep1 seems to be constant during development, the levels of the different Pbx partners changes (Fig. 3.8). For example, the level of Pbx1 decreases while the binding of Prep-Pbx dimers does not change, indicating that the nature of Prep-Pbx complexes present changes during embryonic development. These data suggest that Prep1 could be the limiting factor in heterodimer formation.

The *Hox genes* themselves are spatially expressed while the Prep-Pbx heterodimers are ubiquitously present, suggesting that the effect of Prep on Pbx-Hox probably depends on the DNA target used, and it has the possibility of behaving either as an activator or inhibitor.

Prep1-Pbx complexes are already present in the nucleus of ES cells (data not shown), suggesting that Prep1 could play an important role during the dynamic cellular processes involved in early development. Moreover, the presence of Prep1-Pbx complexes in mouse embryos before the activation of *Hox genes*, suggests a Hox-independent role for these complexes. It is known that Exd and Hth are responsible for antennal development in Drosophila without the involvement of known *Hox genes* (Casares and Mann, 1998). Specific functional experiments are now required to demonstrate the role of Prep-Pbx complexes, both alone and with respect to Hox regulation.

Prep1 and Prep2 represent a new sub-class of Meinox protein family.

The result obtained in collaboration with Fognani suggests that we have cloned a novel member of the Prep family, indicating the existence of two functional groups within the Meinox family of TALE proteins: Prep and Meis

(Fig 3.14). Like Prep1, Prep2 is also able to bind PBC proteins forming dimers which can bind the Prep-Pbx binding sites (Fig 3.14). The complex of Pbx1a with Prep1 formed on O-1 appears to be more robust than one containing Prep2 and Pbx. It is therefore possible that Prep1-Pbx and Prep2-Pbx have different binding sites (Fognani *et al.*, 2002; Haller *et al.*, 2002). This suggests that Prep2 could have a different role from Prep1 in controlling gene expression.

Like Prep1, Prep2 is able to form a ternary complex together with Pbx1 and Hoxb1. Furthermore, the Prep2 ternary complex binds to the PM-PH site sequence present in the *Hoxb2* enhancer (Fognani *et al.*, 2002). However, transfection experiments designed to assess the transcriptional activity of human and murine Prep2 on the *Hoxb1* enhancer revealed a mild repressive effect (Fognani *et al.*, 2002; Haller *et al.*, 2002). Since Prep1 has a positive effect on *Hoxb1* enhancer these data strongly suggest a diversification of activity among the Prep family proteins.

The experiments presented in this chapter includes unpublished data and the data published in the paper: Ferretti, Schulz, H., Talarico, D., Blasi, F. and Berthelsen, J. (1999). "The Pbx-regulating protein Prep1 is present in different Pbx complexed forms in mouse" *Mech. Dev.* **83**: 53-64.
The work on cloning of Prep2 was done in collaboration with Caterina Fognani, DIBIT H.S. Raffaele scientific Institute, Milan. The data are collected in this paper: Fognani, C., Kilstrup-Nielsen, C., Berthelsen, J, Ferretti, E., Zappavigna, V. and Blasi F. (2002). "Characterization of PREP2, a paralog of PREP1, which defines a novel sub-family of the MEINOX TALE homeodomain transcription factors". *Nucleic Acids Res.* **30**:2043-51.

CHAPTER 4
RESULTS AND DISCUSSION II

Segmental expression of Hoxb2 in r4 requires a combined site containing two separated sites that permit cooperative interactions between Prep1, Pbx, and Hoxb1 proteins

Direct auto-and cross-regulatory interactions between *Hox genes* serve to establish and maintain a segmentally restricted pattern of expression in the developing hindbrain. It is known that PBC proteins (Pbx/Exd) are involved in the expression of Hoxb1, which modulates its own expression in rhombomere 4 by forming a heterodimer with PBC factors and binding to specific sites that are present in the autoregulatory element (ARE) of the *Hoxb1* enhancer (Popperl *et al.*, 1995).

Rhombomere r4-specific expression of both Hoxb1 and Hoxb2 depends upon the bipartite *cis Hox* response elements (the PH sites) that bind the group 1 paralogous proteins Hoxa1 and Hoxb1. The DNA-binding ability and selectivity of these proteins depends upon the formation of specific heterodimeric complexes with Pbx proteins. The r4 enhancers from *Hoxb1* and *Hoxb2* have a different number and organization of PH (Pbx/Hox) binding sites, all participating in r4-restricted expression (Popperl *et al.*, 1995; Maconochie *et al.*, 1997). PBC proteins also form stable heterodimers in the nucleus with the Meinox proteins including Prep, Meis and Hth. For this reason, I wanted to investigate if, and in what way Prep1 could modulate Hoxb1 expression. Moreover, since Hox and Meinox proteins interact with independent domains of Pbx, the possibility of ternary complex formation exists.

4.1 Prep1 co-localizes with Pbx and Hoxb1 in murine rhobomere 4.

As shown before, Prep1-Pbx complexes are present in mouse embryos at the same time as Pbx is also interacting with Hox proteins (Fig 3.12).
In previous experiments Jens Berthelsen showed that Prep1-Pbx-Hoxb1 proteins act as co-activators in the transcription of Hoxb1 in *in vitro* transfection experiments (Berthelsen *et al* 1998).

Since Hox proteins display a segmentally restricted pattern of expression in the developing hindbrain, I wanted to verify if Prep1 shows the same localization in the developing hindbrain. Murine embryonic sections at 10.5 d.p.c were stained with either anti-Hoxb1, anti-Prep or anti-Pbx antibodies and PI. Figure 4.1 B shows Prep1 signal in mandibular component of the first branchial arch, hepatic primordium, midgut and lung bud. In addition, Prep1 signal is observed in neuronal tissues, brain, hindbrain, spinal cord and weakly in the heart. Thus, Prep1 protein seems to be widely present in the 10.5 d.p.c. murine embryo. On the contrary, Hoxb1 is only present in r4 as expected (Fig 4.1 A). High magnifications of the hindbrain show the r4-restricted expression of Hoxb1 (Fig 4.1 A, panels D and E) and the diffuse expression of Prep1 (Fig 4.1 A, panels

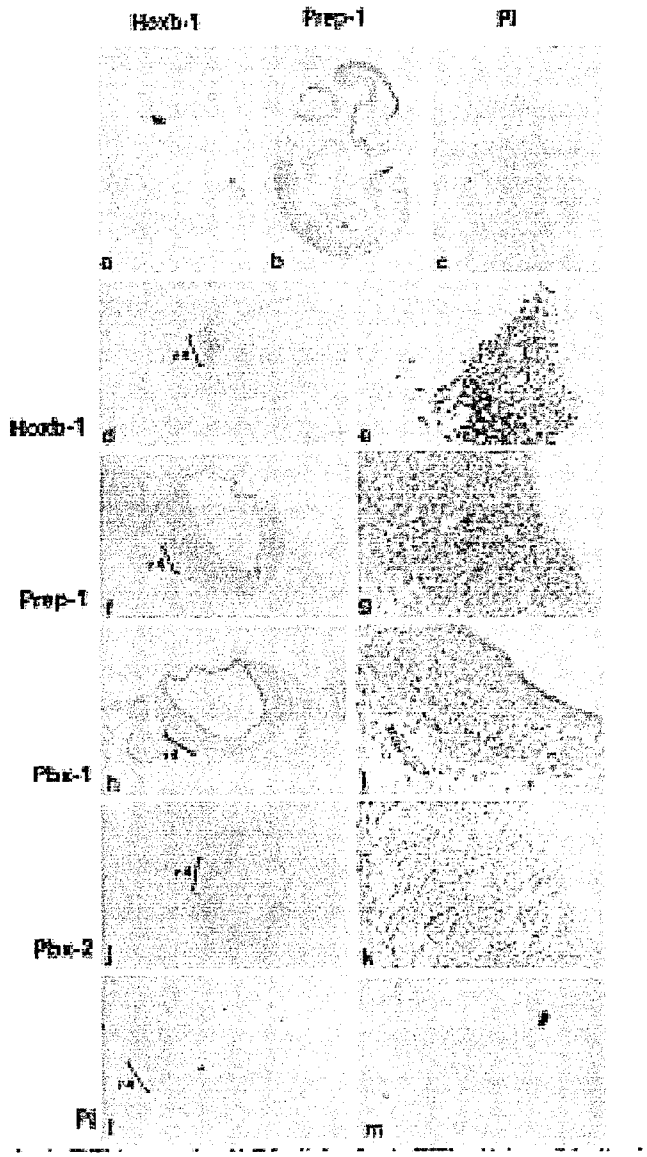


Figure 4.1

Localization of Prep1 protein in mouse embryos. Saggital section of 10.5 d.p.c. mouse embryos were analyzed by immunohistochemical analysis with anti-Prep1 (B, F and G) anti-Hoxb1 (A, D and E), anti-Pbx1 (H and I) , anti-Pbx2 (J and K) and PI (preimmune serum, C, L and M).
Figures D to M: High magnification of contigous sections of the hind brain area.

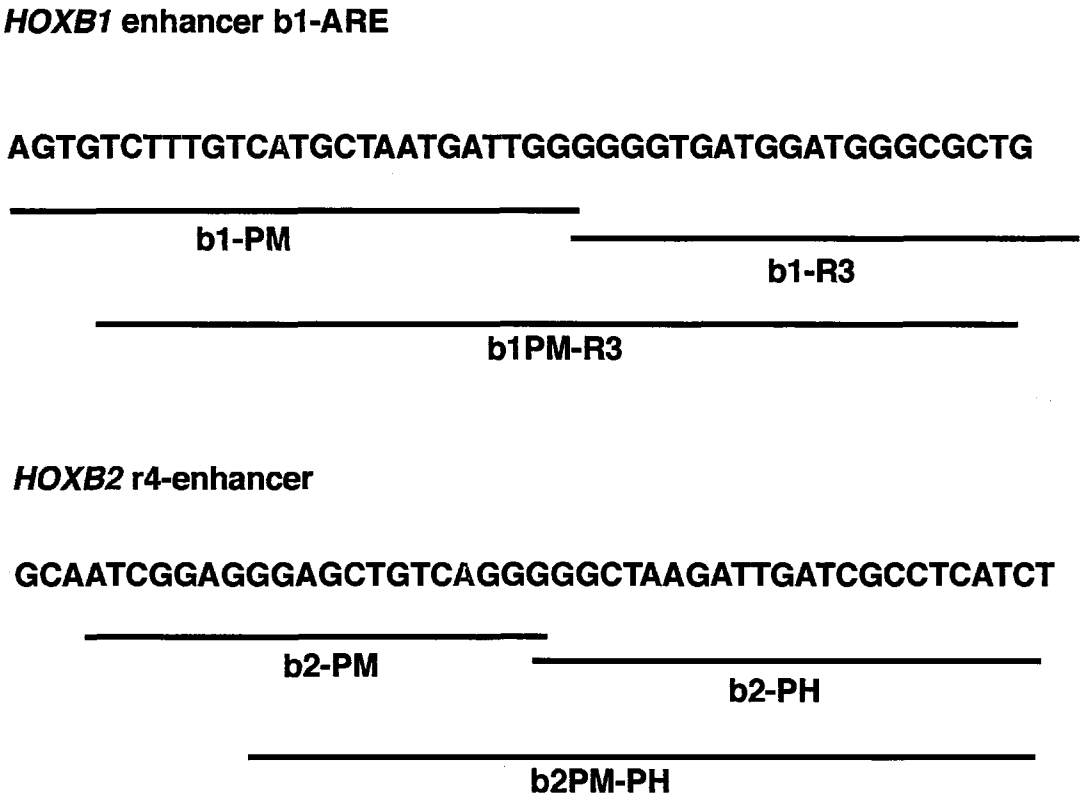


Figure 4.2
The r4-enhancer sequences of *Hoxb1* and *Hoxb2* contain combined PM (Pbx-Meinox binding sites) and PH (Pbx-Hox binding sites) elements. The PH elements are known to promote the specific expression of a reporter gene (LacZ) in r4 of murine embryos (Popper *et al.*, 1995); Macnochie *et al.*, 1997). The sequences of the oligonucleotides used in EMSA are underlined.

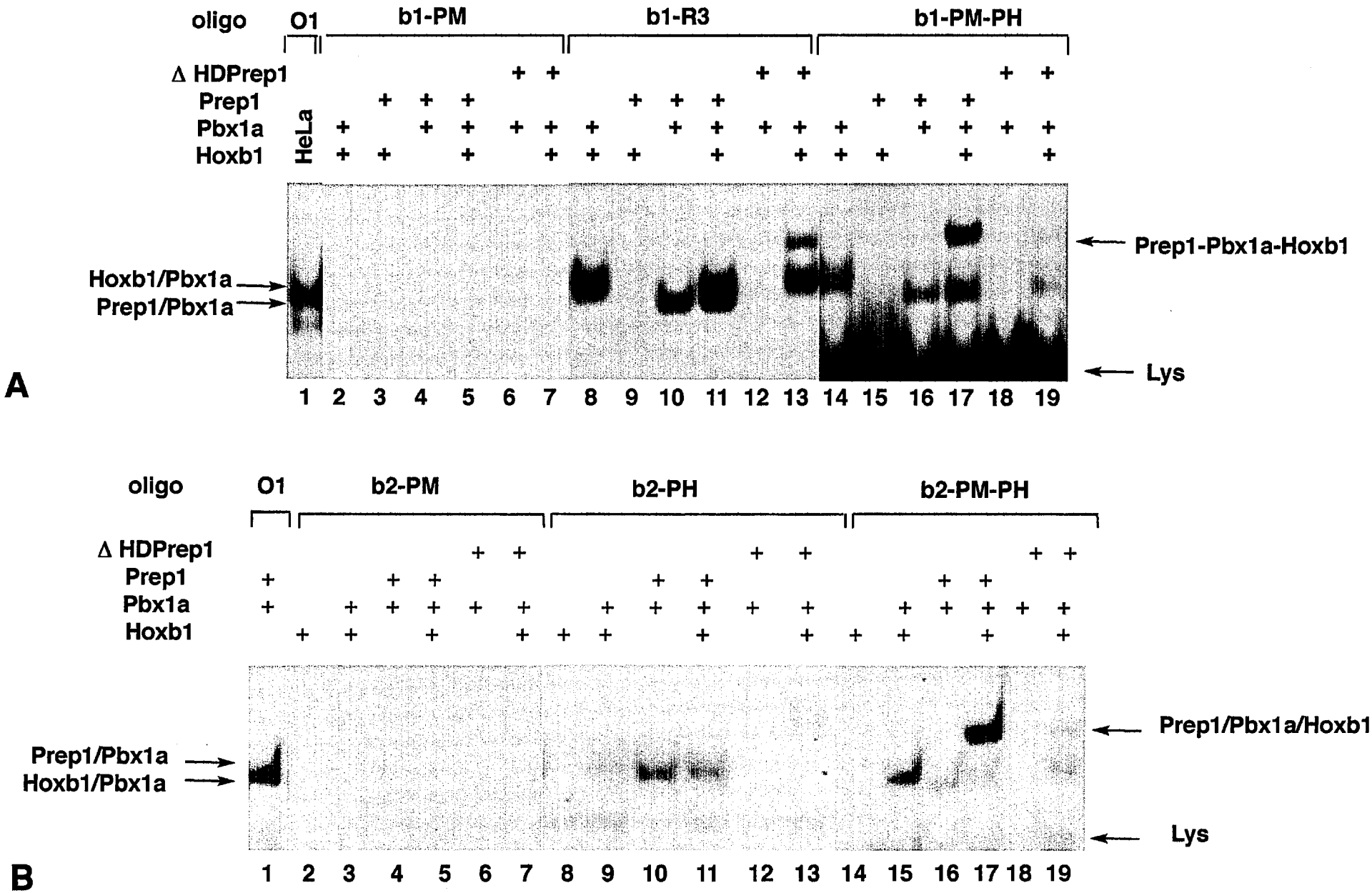


Figure 4.3

EMSA analysis of *in vitro* translated of Prep1, Pbx1a, Hoxb1 and DHDPrep1 with a set of oligos to *r4-Hoxb1* (panel A) and to *4-Hoxb2* (panel B) enhancers(see 4.2 for sequence). A non-specific endogenous activity present in the reticulocyte lysate is marked by Lys. An arrow indicates the slower migrating complex.

F and G), Pbx1 (Fig 4.1 A, panels H and I) and Pbx2 proteins (Fig 4.1 A, panels J and K).

Therefore Prep1, Hoxb1 and Pbx proteins are co-expressed in the r4 at 10.5 d.p.c.. Because of this co-localization, I hypothesized that the formation of a Prep1-Pbx-Hoxb1 ternary complex may occur *in vivo*. This hypothesis is further supported by the fact that: Prep1 and Pbx proteins interact with each other using their N-terminus regions, while Hoxb1 binds to Pbx proteins using the homeodomain regions located at the C-terminus (see Fig 1.19).

4.2 The *Hoxb1* and *Hoxb2* r4-enhancers contain combined Prep1/Pbx and Pbx/ Hox binding sites.

Prep1-Pbx complexes can bind the TGACAG recognition sequences Pbx-Meinox binding site (PM site) but also the Pbx-Hox responsive sequences Pbx-Hox binding site (PH site; Berthelsen *et al.*, 1996/1998, Ferretti *et al* 2000). Since PBC proteins can interact with both Meinox and Hox proteins and thereby affect Hox expression I looked for specific Prep-Pbx (PM) biding sites in *Hox* enhancers.

The examination of the *Hoxb1* and *Hoxb2* enhancer sequences showed the presence of both PH and PM motifs.
In the *Hoxb1* enhancer, in addition to the three PH sites (called R1-R2-R3), a single PM site (TTGTCA) is present 7 bp downstream of R2 and 17 bp upstream of R3 (for detail see Fig 2.2). Also in the *Hoxb2* enhancer I observed a PM site (CTGTCA) located 8 bp upstream of the PH site (Fig 4.2). This site was previously shown to be required for the expression of *Hoxb2* in rhombomere 4 (Maconochie *et al.*, 1997). Next, I examined the *in vitro* binding properties of Prep1, Pbx and Hoxb1 proteins on the putative PM+PH site by EMSA.

4.3 In the *Hoxb1* and *Hoxb2* r4-*enhancers* the *cis* elements PM and PH synergize to form a multimeric complex.

To characterize the molecular interaction between Prep1, Pbx1 and Hoxb1 transcription factors on the combined PM-PH sites of *Hoxb1* and *Hoxb2* enhancers, a set of oligonucleotides were designed and used in *in vitro* binding assays (EMSA; Fig 4.2).

I used oligonucleotides containing either the PM or PH site alone and oligonucleotides containing both sites (PM+PH). As a positive control the original O-1 oligonucleotide was used. The control O1 oligonucleotide containing only the PM site of the *urokinase* enhancer binds the heterodimers Prep1-Pbx1a and Prep1-Pbx1b present in HeLa nuclear extracts (Berthelsen *et al.*, 1996). Labeled oligonucleotides were incubated with different combinations of *in vitro* translated proteins: Pbx1a, Hoxb1, Prep1 and ΔHDPrep1 (a mutated form of Prep1 in which the homeodomain of Prep1 is deleted).

As shown in figure 4.3 A, the b1-PM oligonucleotide, containing the PM site only, is unable to bind any combination of proteins while the b1-R3 oligonucleotide, containing the PH site only, is able to bind the dimeric forms Prep1-Pbx1a and Pbx1a-Hoxb1. Using the b1-R3 oligonucleotide a weak, slower migrating band was observed when the Prep1 homeodomain was deleted

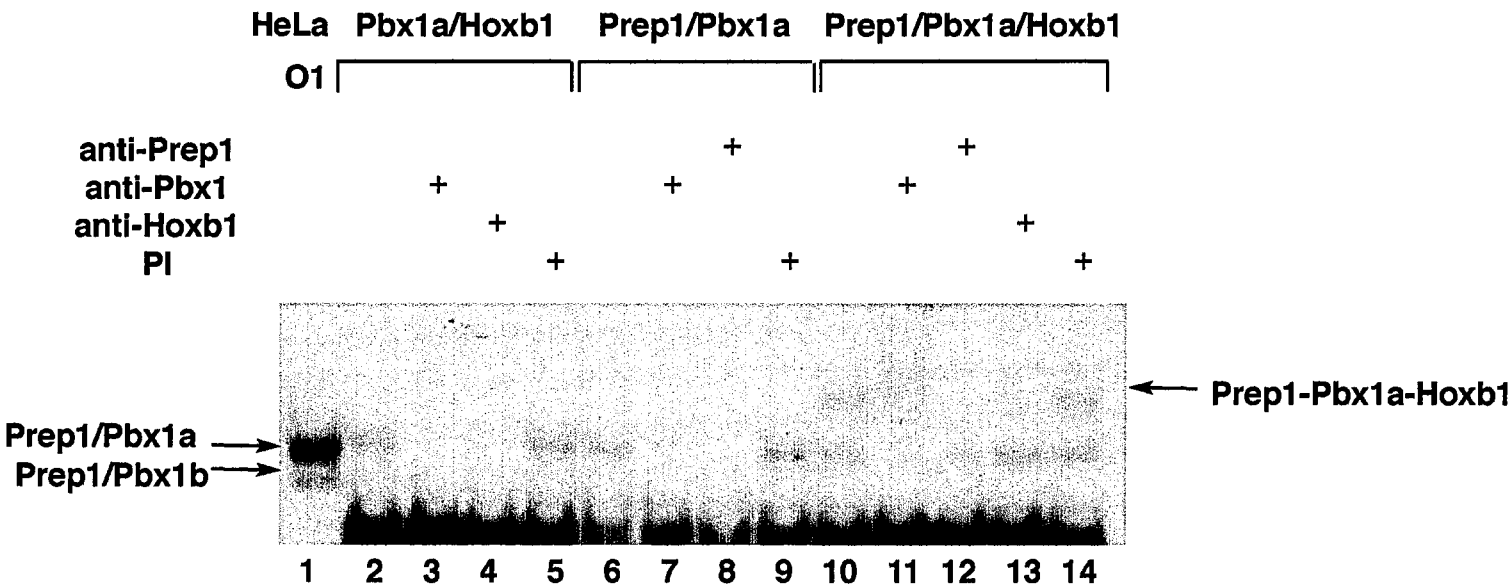


Figure 4.4
The slower migrating DNA-binding activity is a multimeric complex containing Prep1, Pbx1a and Hoxb1. Effect of specific antibodies on the DNA binding activity formed on the b2-PM-PH oligo in the presence of Prep1, Pbx1a and Hoxb1 *in vitro* translated proteins

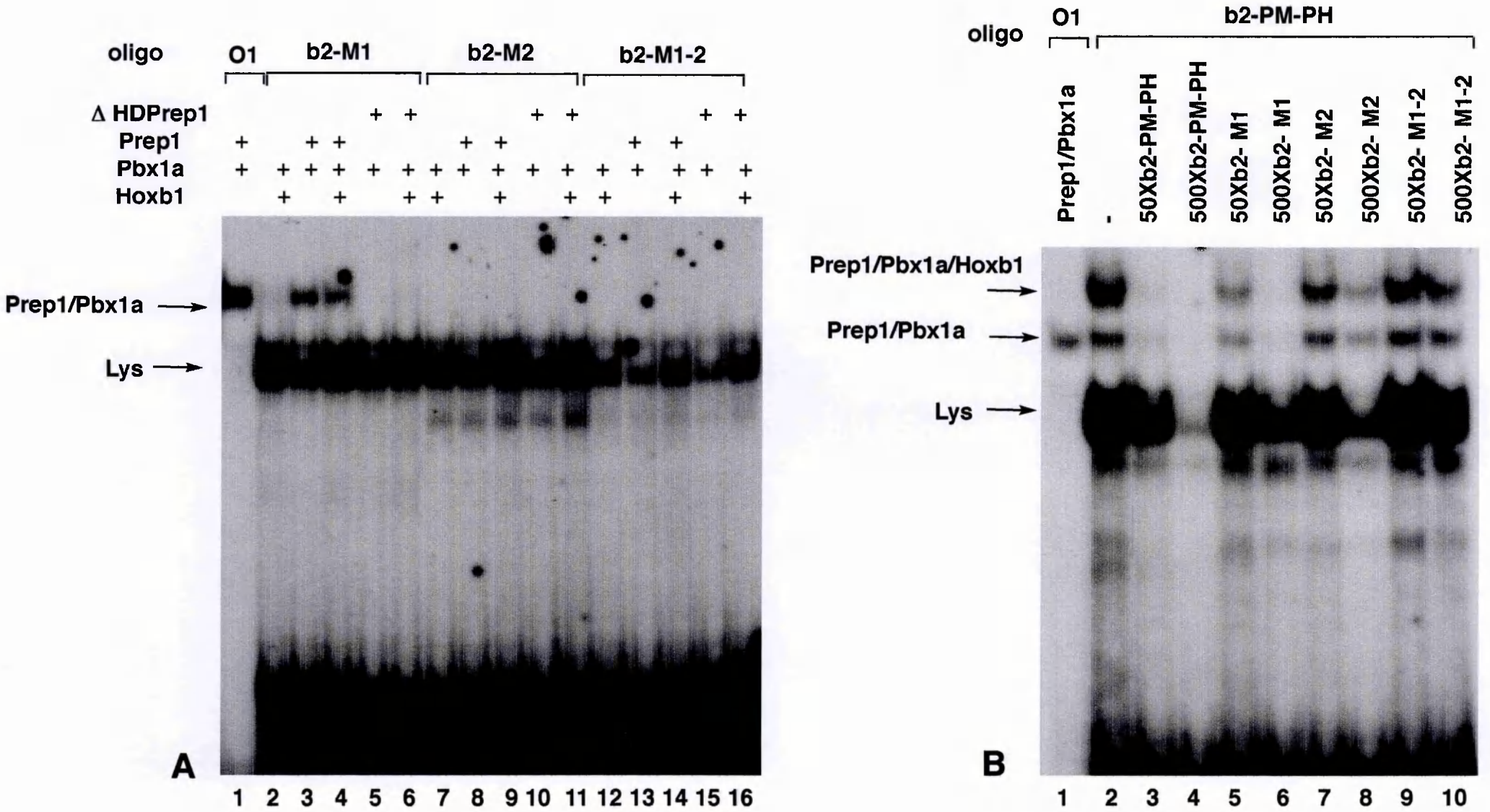


Figure 4.5

PM and PH sites of the *r4-Hoxb2* are both essential for multimeric complex formation. Panel A: Prep1, Pbx1a, Hoxb1 and ΔHDPrep1 were mixed in various combinations with labelled wild type b2-PM-PH oligo as a probe and wild type or mutated b2-PM-PH oligos as cold competitors at 50 and 500 molar excess, as indicated. Lys indicates an endogenous binding activity present in the reticulocyte lysate. The symbol - indicates absence of competitor

(ΔHDPrep1), as already shown by Berthelsen *et al.*, (1998). Moreover, the b1-PM-R3 oligonucleotide, containing both the PM and PH sites, not only bound a Prep1-Pbx1a complex but also formed a slower migrating band possibly corresponding to a multimeric complex. In addition, in the presence of both PM and PH sites, the Prep1 homeodomain played a crucial role in the formation of the multimeric complex. The deletion of Prep1 homeodomain totally prevented the multimeric complex formation (Fig 4.3 A lane 19).

The same results were obtained using oligonucleotides designed from the *Hoxb2* r4-enhancer (Fig 4.3 B). The PM site alone (b2-PM) was not able to bind any combination of proteins, the oligonucleotide containing the PH site (b2-PH) bound the heterodimers Pbx1a-Hoxb1 and Prep1-Pbx1a, while in the presence of both sites (PM+PH) a multimeric complex was formed. Again in the presence of both PM and PH sites the Prep1 homeodomain played a crucial role in the formation of the slower migrating complex (Fig 4.3 B lane 19): its deletion prevented multimeric complex formation.

4.4 The multimeric complex contains Prep1, Pbx, and Hoxb1 proteins.

In order to demonstrate the specificity of the retarded bands, I performed the EMSA experiments again, this time using specific antibodies (Fig 4.4). Both anti-Pbx and anti-Hoxb1 antibodies inhibited the binding of the Pbx1a-Hoxb1 complex, while the Prep1-Pbx1a anti-Pbx-a and anti-Prep1 antibodies inhibited complex. In the presence of all three proteins the formation of dimeric bands was partially inhibited by anti-Prep1 or anti-Hoxb1 antibodies, suggesting that the dimeric bands are composed of both Prep1-Pbx1a and Pbx1a-Hoxb1 heterodimers. The slower migrating complex was completely inhibited by anti-Prep1 or anti-Hoxb1 antibodies. The anti-Pbx1 antibodies completely abolished the binding of the dimeric form and partially inhibited the slower migrating band. I therefore conclude that the slower migrating band contains Pbx1a, Prep1 and Hoxb1.

4.5 Both PM and PH sites are required for the ternary complex formation.

I next investigated the co-operation between PM and PH sites in the formation of the slower migrating band.

EMSAs were repeated using oligonucleotides that were point mutated in either the PM, PH or both sites (Fig 4.5 A). Point mutations in the PM site (b2-M1) inhibited the formation of the slower migrating band leaving the binding of the heterodimeric forms Prep1-Pbx1a and Pbx1a-Hoxb1 to the PH site unaffected. Point mutations in the PH site (b2-M2) abolished the formation of both the slower migrating band and of the heterodimeric forms Prep1-Pbx1a and Pbx1a-Hoxb1. Again the PM binding site alone was not able to bind any combination of proteins, suggesting that this site is a low affinity binding-site for Prep1-Pbx complexes. The double mutant (b2 M1-2) was unable to bind any complexes as expected.

According to the data described above I conclude that the PH site alone is sufficient for heterodimer formation, but that both sites are required for the ternary complex formation. In fact, mutation of only one of the two sites was necessary to block the formation of high molecular weight complexes.

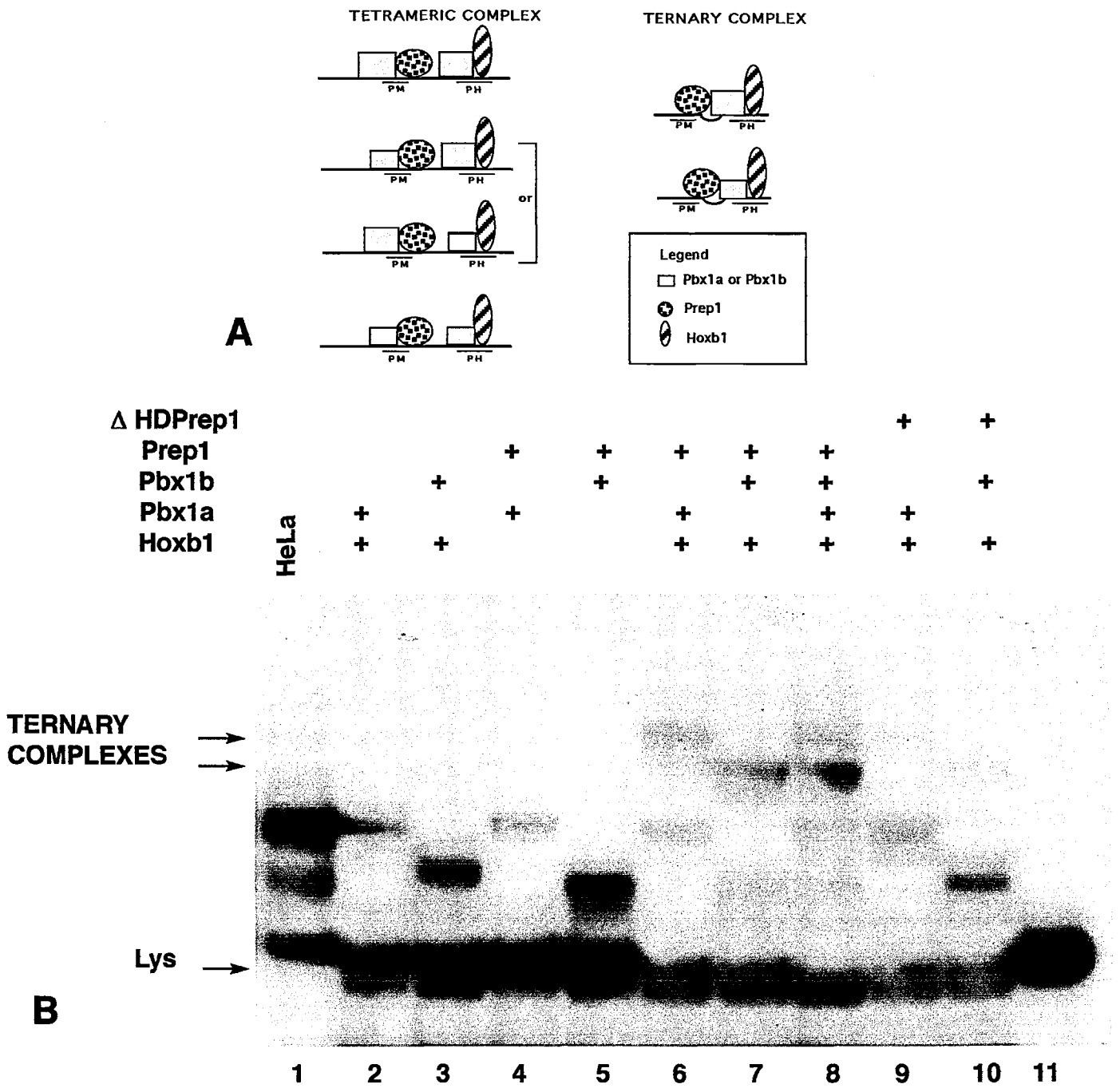


Figure 4.6

The slower migrating bands are ternary complexes composed of Prep1, Pbx1 and Hoxb1 proteins: panel A: scheme of possible complexes binding the b2-PM-PH oligo. The slower migrating multimeric complex could be either a Prep1-Pbx1-Hoxb1 ternary complex (right) or two heterodimers, Prep1-Pbx1 and Pbx1-Hoxb1, bound to the same DNA molecule (left). A mixture of *in vitro* translated Prep1, Pbx1a, Pbx1b and Hoxb1 proteins was used to distinguish between ternary and tetrameric complexes. Panel B: EMSA performed with different combinations of *in vitro* translated proteins and b2-PM-PH labeled oligo (the protein compositions of each binding reaction is shown on the top). Arrows indicate the migration of the two ternary complexes. HeLa nuclear extract show Prep1-Pbx1a and Prep1-Pbx1b complexes, while Lys corresponds to activity present in reticulocyte lysate.

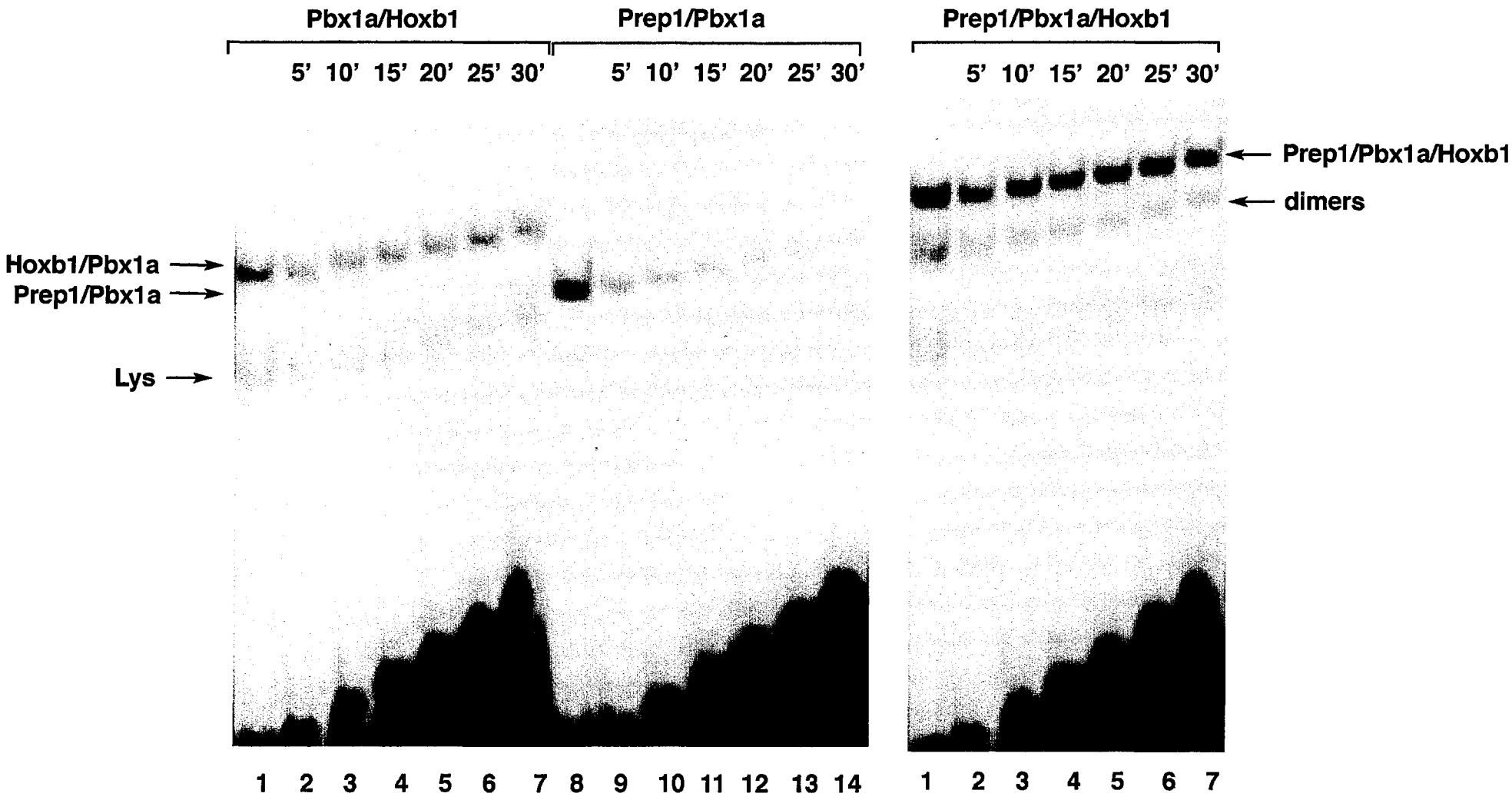


Figure 4.7

Analysis of the dissociation rate of the complexes formed on b2-PM-PH; the oligo was incubated for 10 minutes on ice with *in vitro* translated Prep1, Pbx1a ,Pbx1b and Hoxb1 protein mixtures, as indicated on top. A 100 fold excess of unlabelled b2-PM-PH oligo was added and aliquots analyzed by EMSA at different times, as indicated. Endogenous activity present in reticulocyte lysate is indicated by Lys.

These experiments were further confirmed by competition assays. EMSAs in figure 4.5 B show that a mutated PM site (b2-M1) still competes with the wild type site for binding to the heterodimer and the ternary complex, although less efficiently than the wild type site (b2-PM-PH). The oligonucleotide mutated in the PH site (b2-M2) showed only partial competition, even in the presence of a 500-fold molar excess. The oligonucleotide containing the double mutation (b2-M1-2) no longer competed for binding. These data, and that the PH site represents the major binding site for the dimeric complexes while the PM and PH sites synergize in multimeric complex formation.

4.6 The slower migrating complex formed on the combined PM and PH sites is a ternary Prep1-Pbx1a-Hoxb1 complex and not a tetrameric Prep1-Pbx1 and Pbx1-Hoxb1 complexes.

The migration properties of the multimeric complex bound to the combined PH-PM site are compatible with the binding of either two dimeric complexes or of a ternary Prep1-Pbx-Hoxb1 complex. In order to investigate if the slower migrating band was a ternary complex or a double dimer in which the binding of the first dimer facilitated the binding of the second one I repeated EMSA experiments in the presence of two different Pbx splice variants, Pbx1a and Pbx1b, having a different C-terminal extension (a schematic representation of the experiments is shown in figure 4.6). Both splice variants bind Prep1, Hoxb1 and the target DNA sequence, forming heterodimers of different molecular weights.

If, in the presence of all four proteins, (Prep1, Hoxb1, Pbx1a and Pbx1b) two dimers are formed (Prep1-Pbx1 and Pbx1-Hoxb1), the EMSA should result in three high molecular weight bands, corresponding to Prep1-Pbx1a, Prep1-Pbx1b and Hoxb1. However, in the case of ternary complex formation I would expect to see only two bands only, corresponding to Prep1-Pbx1a-Hoxb1 and Prep-Pbx1b-Hoxb1. As shown in Figure 4.6, in addition to the two dimeric forms Prep1/Pbx1b (lane 5) and Prep1-Pbx1a (line 3), I also observed a slower migrating band in the presence of all four proteins (lane 7). This binding activity is showing only two slower migrating bands. I therefore conclude that the slower migrating band is a Prep1-Pbx-Hoxb1 ternary complex and not two dimeric forms.

4.7 The Prep1-Pbx-Hoxb1 ternary complex has a higher stability then the Prep1-Pbx dimer.

In order to compare the stability of the ternary complex (Prep1-Pbx1a-Hoxb1) versus the heterodimeric forms (Pbx1a-Hoxb1 and Prep1-Pbx1a) I analyzed the dissociation rate of these complexes formed on the b2-PM-PH oligonucleotide (Fig. 4.7). Proteins were incubated with labeled b2-PM-PH for 10 minutes and then a 100-fold excess of the same unlabeled probe was added as competitor. Samples were withdrawn every 5 minutes and analyzed by EMSA. Figure 4.7 shows how the Prep1-Pbx1a complexes dissociate from the b2-PM-PH oligonucleotide with a half-life of about 20 minutes, while Pbx1-Hoxb1 and the ternary complex are much more stable. These data show that the ternary complex Prep1-Pbx1a-Hoxb1 is more stable than the dimer Prep1-Pbx1a. Since

Prep1-Pbx complexes are formed in the absence of DNA, these data suggest that the subsequent binding of Hoxb1 to the Prep1-Pbx complexes confers higher stability to the ternary complex.

4.8 Nuclear extract of P19 cells contain constitutive Pbx-Prep1 and are induced by retinoic acid to form Prep1-Pbx-Hoxb1 complexes.

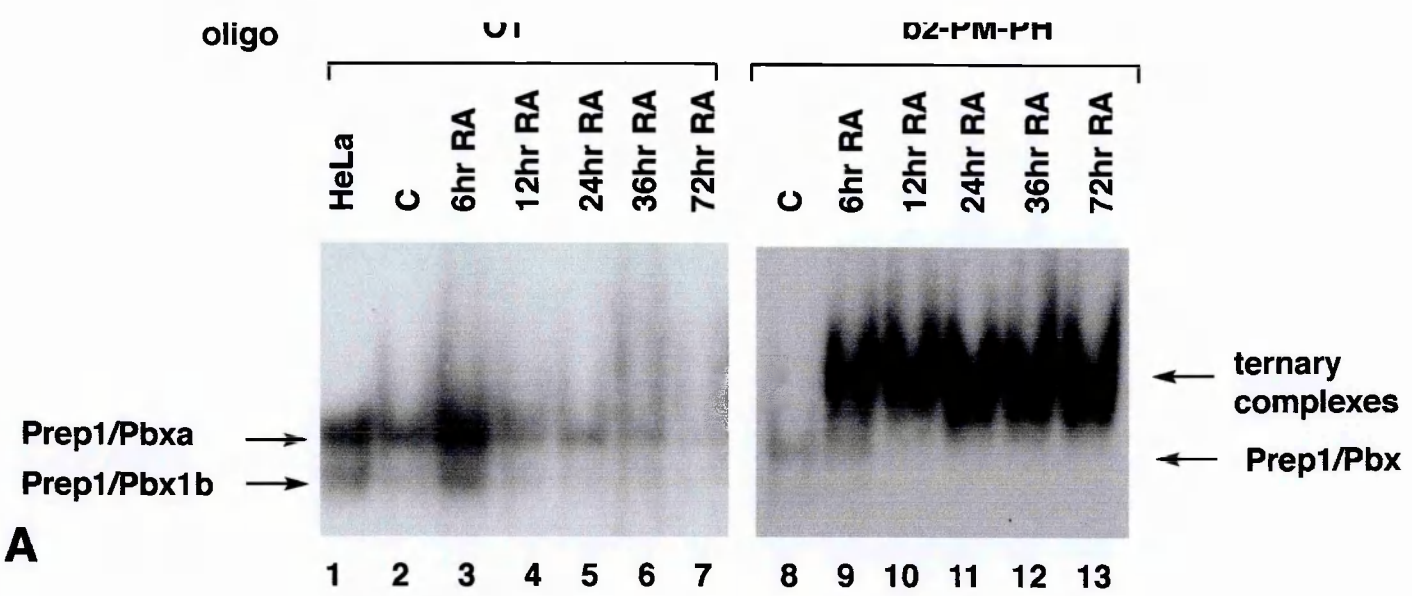
In cultured embryonic carcinoma cells, treatment with retinoid acid (RA) induces the expression of *Hox genes* (Simeone *et al.*, 1990; Marshall *et al.*, 1992) and activation of Pbx proteins (Knoepfler and Kamps., 1997).

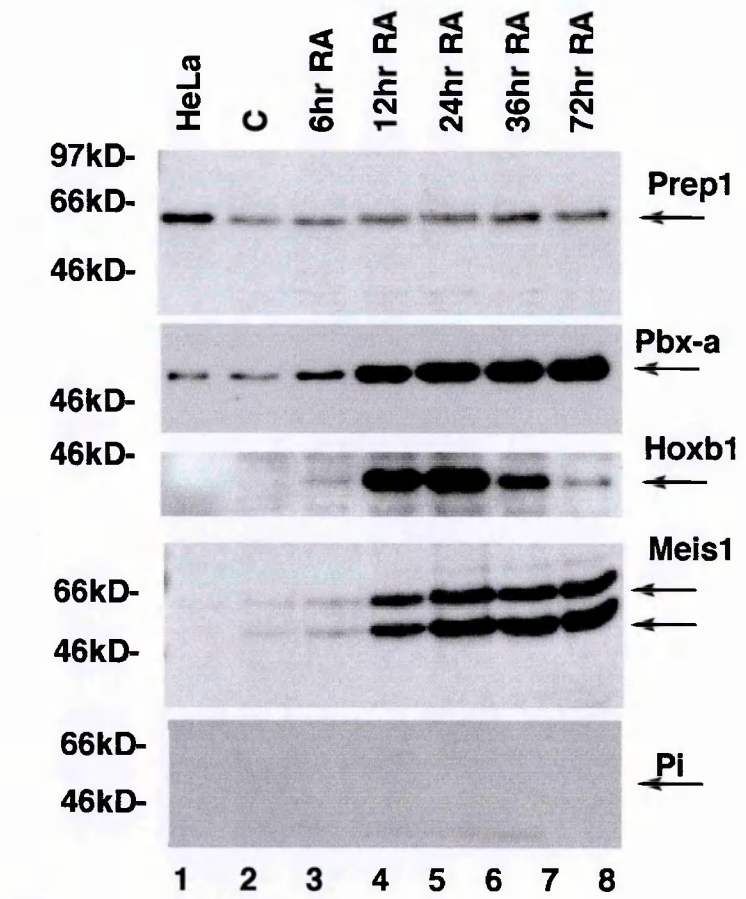
In order to test the formation of ternary complexes *in vivo*, I analyzed the retarded bands formed after incubation of the b2-PM-PH oligonucleotide with nuclear extracts obtained from untreated and RA-induced P19 cells by EMSA. As shown in figure 4.8 A, nuclear extract from untreated P19 cells incubated with labeled O1 oligonucleotide produces two retarded bands co-migrating with UC and LC from HeLa extract and corresponding to Prep1-Pbx1a and Prep1-Pbx1b heterodimers. 6 hours after induction with 10^{-5} mM RA, the presence Prep1-Pbx dimers diminished and was substituted by a slower migrating band comigrating with the ternary complexes formed with *in vitro* translated Prep1-Pbx-Hoxb1 proteins.

I also showed P19 nuclear extract after RA induction by immunoblotting. As shown in figure 4.8 B, Prep1 was present in control extracts and slightly increased upon RA induction. On the other hand, reactivity to anti-Pbx-a antibodies was present in untreated extract and was strongly enhanced by RA induction (Fig 4.8 B). Pbx antibodies were specific for the 50kDa isoform only. The anti-Meis1 antibody (Meis1a/1b) detected two very weak bands under basal condition, which were dramatically increased 12 hours after RA treatment (Fig 4.8 B). While Hoxb1 was absent in untreated cells, it was present at 6 hr and strongly induced at 12 hr decreasing there after (Fig 4.8, panel B). Binding activity of the 6 hr induced P19 nuclear extracts in the presence of labeled b2-PM-PH was analyzed by competition with unlabeled wild-type or mutated oligonucleotides (Fig 4.8 C). Wild-type b2-PM-PH totally competed with labeled b2-PM-PH at a 50 fold excess, while mutant PM (b2-M1) competed well for Prep1-Pbx binding at 50 excess but only weakly for the slower migrating complex, binding at a 500-fold excess. Mutant PH (b2-M2) weakly competed for the binding to both dimeric and multimeric forms but only at 500-fold excess, while the double mutant (b2-M1-2) did not compete at all. Thus, nuclear complexes present in induced P19 cell reproduced the sequence specificity towards the b2-PM-PH oligonucleotide already shown with *in vitro*-translated proteins (Fig 4.4 and Fig 4.5).

The protein composition of all the complexes forming on the b2-PM-PH oligonucleotide was tested using specific antibodies. The UC and LC that formed in the absence of RA treatment were completely inhibited by anti-Prep1 and anti-Pbx-a antibodies, while the anti-Pbx1 antibody only partially inhibited UC and LC formation, indicating the presence of Prep1 and several Pbx isoforms in the two complexes (Fig 4.9).

The formation of slower migrating band after 6hr with RA was partially, but specifically, inhibited by anti-Prep1, anti-Pbx or anti-Hoxb1 antibodies. However, the multimeric complexes were completely inhibited by co-incubation





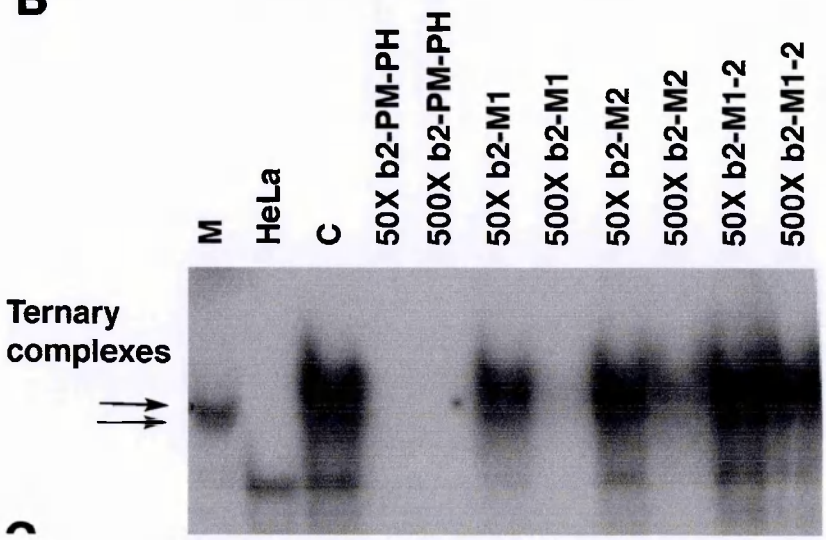


Figure 4.8

Time course analysis of nuclear extracts from control and RA induced p19 cells. HeLa represents the HeLa nuclear extract containing the Prep1-Pbx1a and Prep1-Pbx1b dimers. Nuclear extracts purified from untreated cells are indicated by C.

Panel A: EMSA with labeled O-1 and 2-PM-PH oligos. A slower migrating band able to bind b2-PM-PH but not O-1, appeared after RA treatment.

Panel B: immunoblotting analysis using anti-Prep1, anti Pbxa, anti-Hoxb1, anti-Meis1 or PI antibodies.

Panel C: Competition assay using labeled b2-PM-PH and nuclear extract after 6 hours RA treatment. In the binding-competition experiments labeled b2-PM-PH and unlabeled wild type and mutated (50- and 500-fold excess) b2-PM-PH oligonucleotide were used .

The RA-induced slower migrating band has the same binding specificity, as *in vitro* translated Prep1-Pbx1a-Hoxb1 ternary complex.

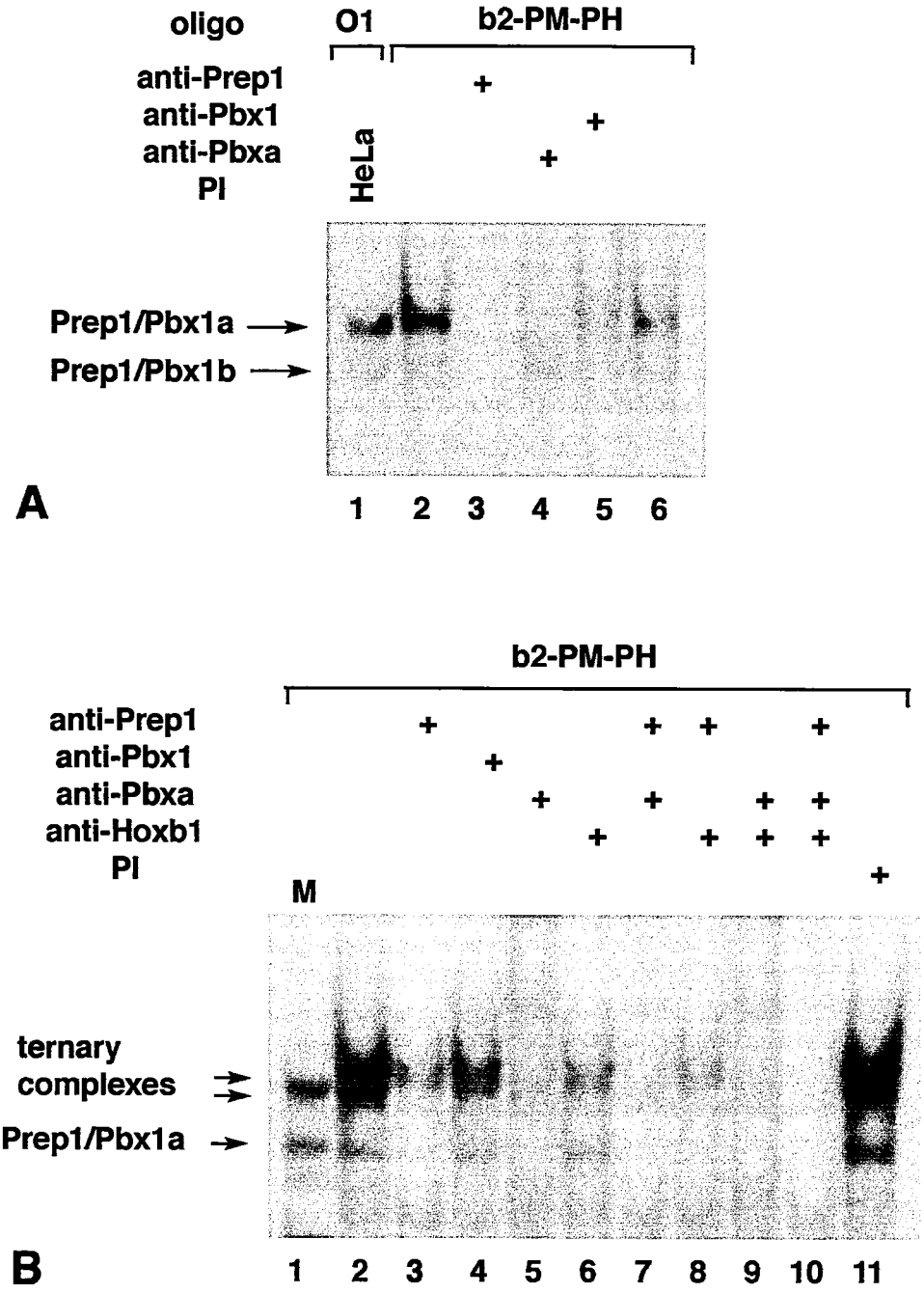


Figure 4.9

EMSA using labeled b2-PM-PH oligo and P19 nuclear extract in the presence of specific antibodies.

Panel A: Anti-Prep1 and anti-Pbx-a antibodies compete for the binding of the complexes formed in the untreated cells extracts.

Panel B: The combination of anti-Prep1, anti-Pbx-a and anti-Hoxb1 inhibits the binding of all complexes formed after RA treatment.

with all three antibodies. These different quantitative effects suggest a differential exposure of the various epitopes in the heterodimeric complex. We can conclude that in P19 cells the slower migrating DNA-binding complex that appears on Hoxb1 is induced with RA. P19 cells contain different members of the Pbx family which could all participate in the complex formation. In addition to Pbx other Meinox proteins could form the multimeric complex.

According to our data 6hr after RA-induction, the ternary complex contains Pbx1a, Hoxb1 and Prep1. However we cannot exclude that at later times Meis substitutes or supplements Prep1.

4.10 Characterization of PM binding site affinity

In figure 4.3 I showed that the PM motif in r4-*Hoxb2* and r-4 *Hoxb1* enhancers, which contains the canonical binding site for Prep1-Pbx complexes is not capable of binding Prep-Pbx dimers. Other PM sites, for example that of the human u*rokinase, interlukin 3* and *stromelysin* enhancer, easily bind Prep1-Pbx dimers as determined by methylation interference analysis (Berthelsen *et al.*, 1996). Therefore, the PM sites in *Hoxb*1 and *Hoxb2* enhancers seem to be low affinity binding sites for Prep1-Pbx heterodimers. However, when located near a PH site, they permit the formation of Prep1-Pbx-Hoxb1 ternary complex (Ferretti et el.,2000). Possibly, the binding of Prep1-Pbx dimer to the *urokinase* PM site is facilitated by adjoining bases present in the *urokinase* and missing in the *Hoxb1* and *Hoxb2* enhancers.

Is the low affinity for Prep1-Pbx required for ternary complex formation? To answer this question I first compared the sequence flanking the PM site in the *urokinase* enhancer (O-1 oligonucleotide) and in the *Hoxb2* enhancers (b2-PM oligonucleotide).
I observed that the three bases upstream and the three bases downstream immediately flanking the TGACAG core sequence were different (in the two enhancers: CTT versus GAG at the 5′ and CTC versus GGG at 3′ (Fig 4.10),
Starting from these observations, I generated a set of mutated oligonucleotides (Z1 to Z6; Fig 4.10) to study the role played by those triplet in determining the site binding affinity. Using b2-PM as back-bone sequence I introduced single or multiple substitutions upstream and downstream of the PM-core sequence, according to the O1 sequence (Fig 4.10).

EMSA were performed using *in vitro* translated proteins and labeled O1. Unlabeled wild-type b2-PM and mutated oligonucleotides were used as competitors (Fig 4.11). As shown in Figure 4.11 competition I have tested a set of oligonucleotide containing single or multiple substitutions upstream and downstream of the PM-core binding site.

I found that all Z oligonucleotides in which the 3′-sequence GGG was substituted with the 3′ sequence CTC from O1, were able to compete for the binding of Prep-Pbx dimers to O1 (Fig 4.11). In order to define a minimal sequence required to assemble the Prep1-Pbx dimers only one (C in Z5) of two (CT in Z6) bases were substituted (Fig 4.11). The results show that all three 3′ CTC bases present in O1 are required to convert a low affinity PM site into a high affinity one (Fig 4.11).

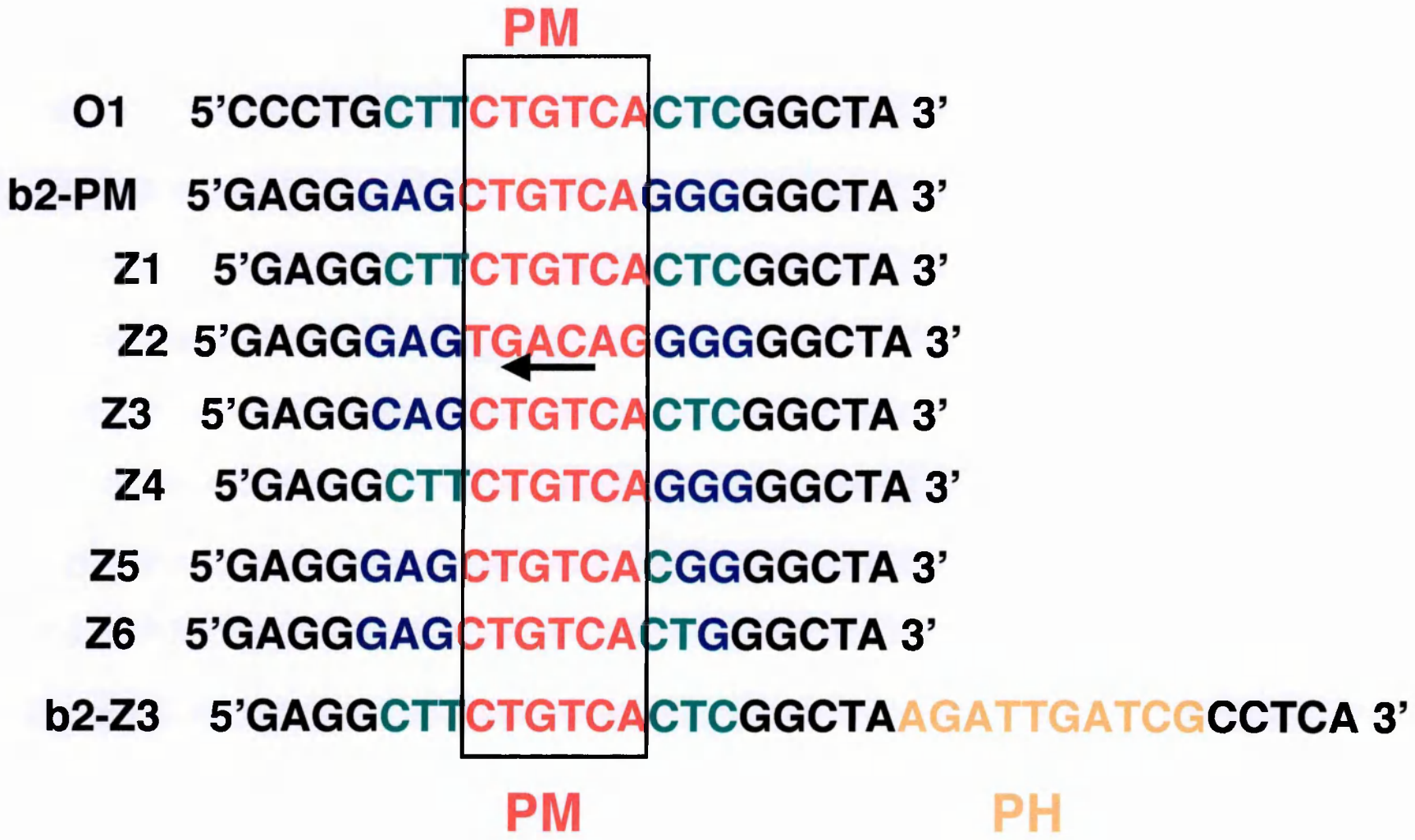


Figure 4.10
Schematic representation of the high affinity O-1 low affinity (b2-PM) and mutated (Z set) PM sites. In the Z set of oligonucleotides, modifications were introduced into the b2-PM sequence. The arrow indicated the mutation in which the PM site sequence was inverted.

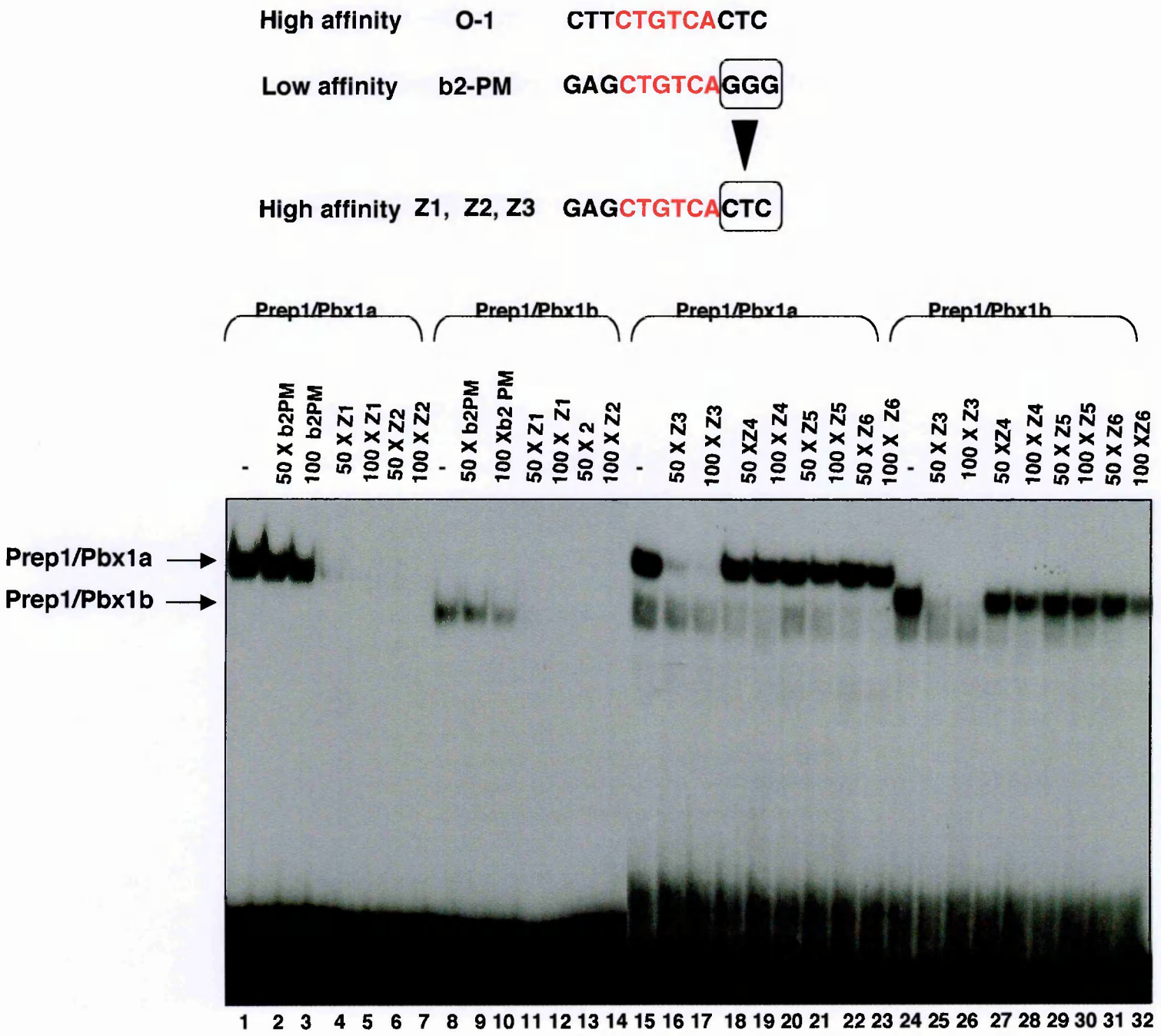


Figure 4.11
Analysis of PM sites binding activity affinity. Competition assays in which the binding of *in vitro* translated Prep1-Pbxa and prep1-Pbx1b dimers to labeled O-1 was competed with 50-or 100 fold excess of unlabeled wild type (b2-PM) or mutated (2 set) oligonucleotides. The CTC sequence present downstream of PM in O-1, converts the low affinity PM site in b2-PM into a high affinity binding site (lanes 21, 22 and 23). The symbol - indicated the absence of any competitor.

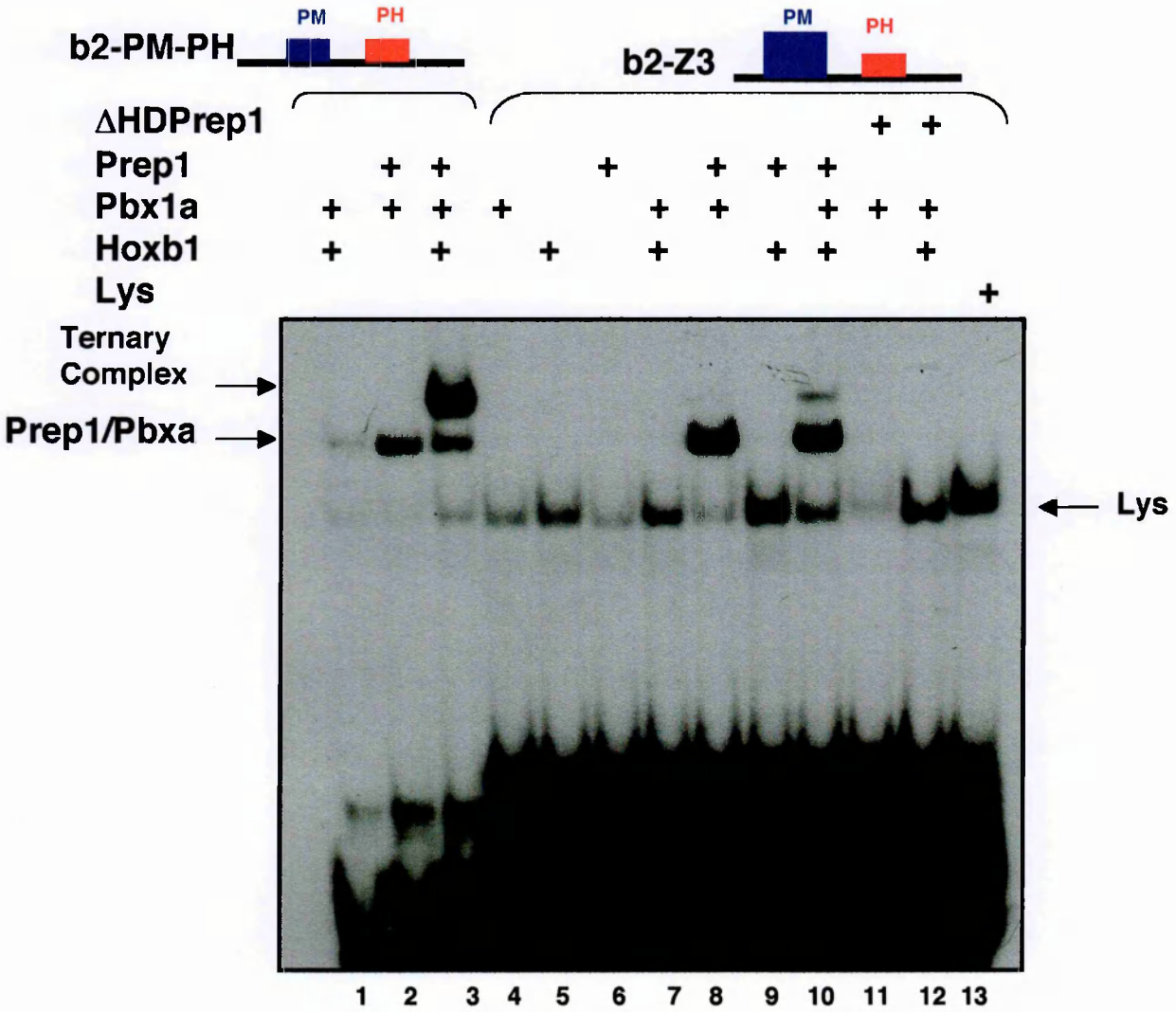


Figure 4.12
EMSA analysis of Prep1, Pbx1a and Hoxb1 in vitro translated proteins and b2-Z3 oligonucleotide, in which the low affinity PM site was substitute with a high affinity site. This substitution interfere with the Pbx/Hoxb1 binding (lane 7) and thus with the ternary complex formation (lane 10).

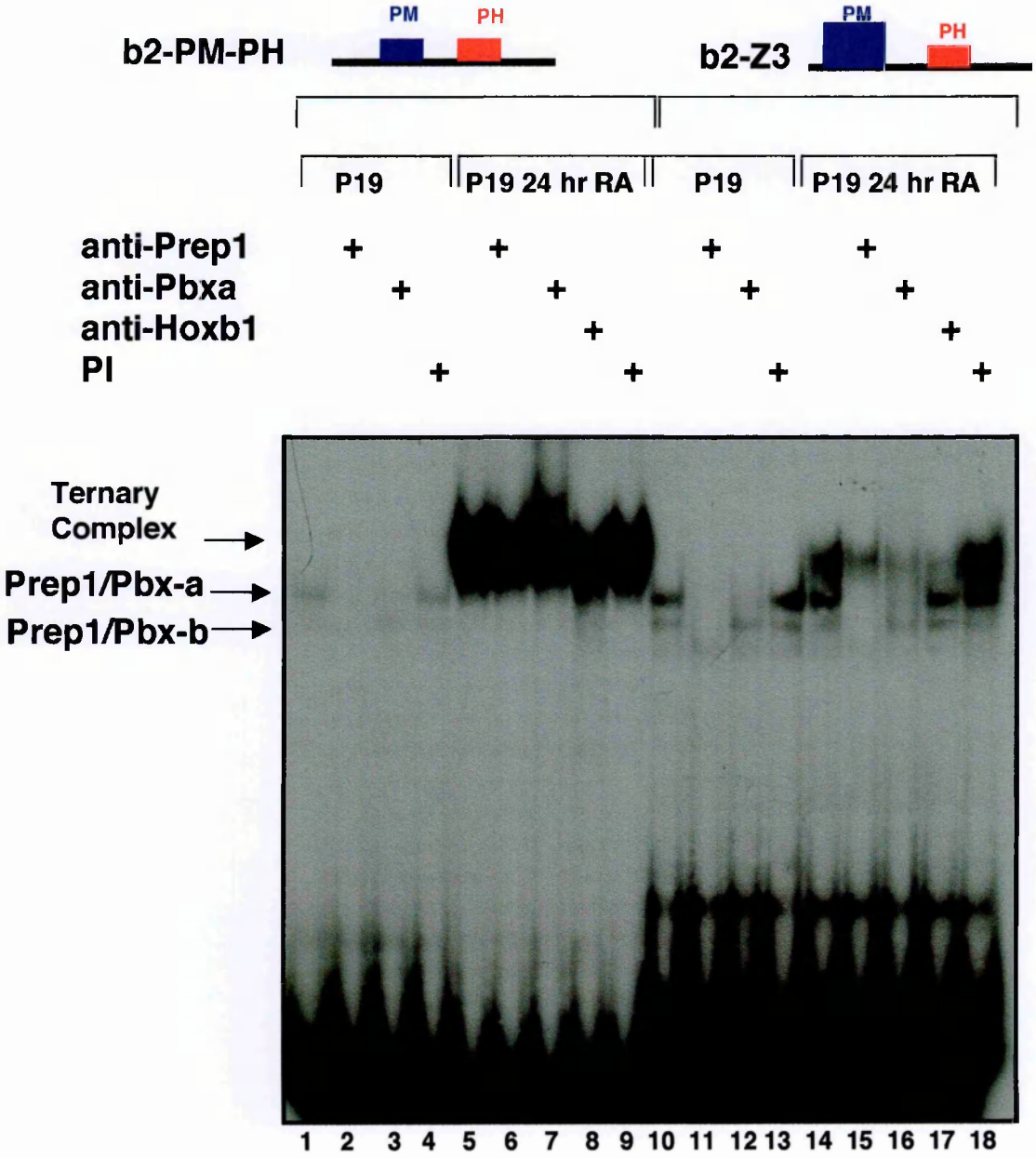


Figure 4.13
EMSA analysis of nuclear extracts purified from untreated and RA-treated P19 cells. b2-Z3 oligonucleotide containing the high affinity PM site strongly binds Prep1-Pbx-a and Prep1-Pbx-b dimers (compare lanes 1 and 10). The presence of high affinity PM site inhibits the assembly of ternary complex obtained with RA treated P19 nuclear extracts (lane 14).

To test if the presence of high affinity PM site can alter the ternary complex formation, I designed an oligonucleotide (b2-Z3) containing both a PM and a PH site required for the ternary complex binding, but in which the low affinity PM site was replaced by a high affinity one Z3 (Fig 4.12).
b2-Z3 was compared with the wild-type b2-PM-PH oligonucleotides for its ability to form a ternary complex in the presence of *in vitro* translated Prep1, Pbx1a and Hoxb1 proteins.

As shown in figure 4.12 Prep1-Pbx1a heterodimers bind the b2-Z3 with higher affinity than to the wild type b2-PM-PH oligonucleotide (Fig 4.12; lane 8). However, this oligonucleotide is not able to form the ternary complex (Fig 4.12 lane 10), confirming the hypothesis that a high affinity PM site interferes with the formation of a ternary complex, probably because of the preferential binding of Prep-Pbx dimers creates a steric hindrance that hampers the binding of Pbx-Hox dimers (lane 7) and thus the ternary complex formation (lane 10). I also tested the binding properties of the b2-Z3 using the nuclear extract obtained from either untreated or RA treated P19 cells (Fig 4.13). Figure 4.12 shows that the b2-Z3 bindsPrep1-Pbx heterodimers both in untreated and in RA-treated cells with higher affinity than the wild type (b2-PM-PH). After RA-induction it also forms the ternary complex but with much less efficiency than wild-type oligonucleotide b2-PM-PH (Fig 4.13).

Similar results were obtained analyzing the PM site of *Hoxb1* enhancer (data not shown).

4.11 The PM site in the *Hoxb2*, but not in *Hoxb1*, enhancer is required for *in vivo* r4 restricted expression.

To determine the role of PM sites in the *Hoxb1* and *Hoxb2* enhancer, *in vivo*, transgenic mouse embryos were generated carrying the *Hoxb1* or *Hoxb2* wt or mutated enhancers fused to the reporter LacZ gene (in collaboration with Robb Krumlauf lab). The LacZ reporter activity was analyzed in transgenic embryos at 9.5 d.p.c. The 181 bp StuI-fragment of the *Hoxb2* enhancer and the β-actin promoter are able to drive the r4 restricted expression of the transgene. This sequence contains a bipartite Pbx/Hox binding site at position 97 that is essential for enhancer activity (Maconochie *et al.*, 1997). The deletion of the first 85 bp (containing the PM motif) completely abolished LacZ expression in r4 (Fig 4.14B). Transgenic embryos carrying mutations in the PM site were also generated. The PM mutation prevents the binding of Prep1/Pbx complexes. As shown in figure 4.10C, the wild-type enhancer drives the expression of the reporter gene in r3, r4, r5 and in more posterior rhombomers. Mutations in the PH site specifically abolished the reporter gene expression in r4 (Fig 4.14D), while the mutation in the PM site had an even stronger influence (Fig 4.14E),. The expression of the transgene was abolished not only in r4 but also in the posterior rhombomers (Fig 4.14E), leaving the expression in r3 and r5, which are known to be under the control of a different transcription factor called Krox20 (Sham *et al.*, 1993; Vesque *et al.*, 1996).

Thus, the presence of both PM and PH sites are essential for r4 restricted expression, suggesting an essential role for the ternary complex Prep1-Pbx-Hoxb1 in Hoxb2 expression. Furthermore, it appears that the PM site has additional influences on Hoxb2 expression which are independent of the PH site.

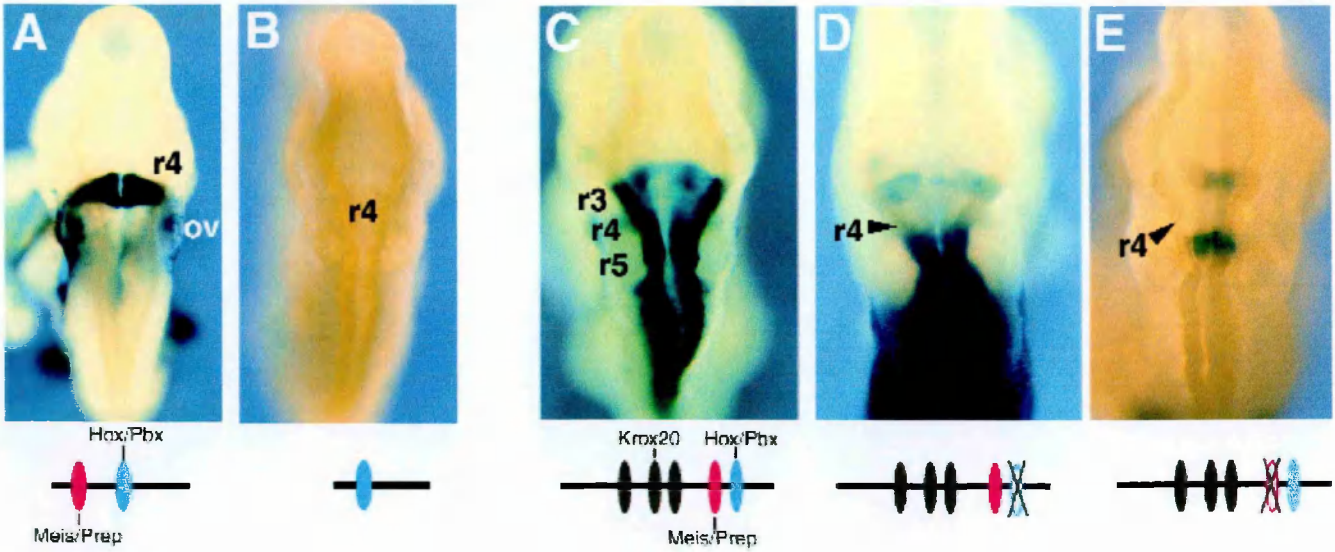


Figure 4.14

PM cooperates with PH for r4-restricted expression of the *Hoxb2* enhancer. Dorsal views of LacZ transgene expression in the hind brain of 9.5-9.75 d.p.c. mouse embryos.
Blue: the bipartite PH sites; purple: PM site; gray: the three Krox20 binding sites. An X indicated the presence of mutations.
A) Strong r4-restricted expression of LacZ fused to wild-type 181 bp StuI enhancer. B) Loss of r4-restricted expression upon deletion of the first 85 bp, leaving the PH site intact. C) Reporter staining in r3, r4, r5 and the posterior region of the neural tube and mesoderm mediated by a wild-type 2.1 Kb bamHI-EcoRI *Hoxb2* enhancer. D) Point mutation in the PH site specifically eleminates LacZ expression in r4 (arrowhead). The lower level of r3 staining in this and in (E) is due to the embryos being slightly older than C when Krox20 dependent expression as already begun to decrease. E) Mutations in the PM site result in the loss of transgene expression in r4 (arrowhed) and also in more oposterior regions. The absence of more posterior expression, not seen with the mutation in PH site, indicates that the PM site may additional roles in *Hoxb2* expression. OV: Otic Vecicle

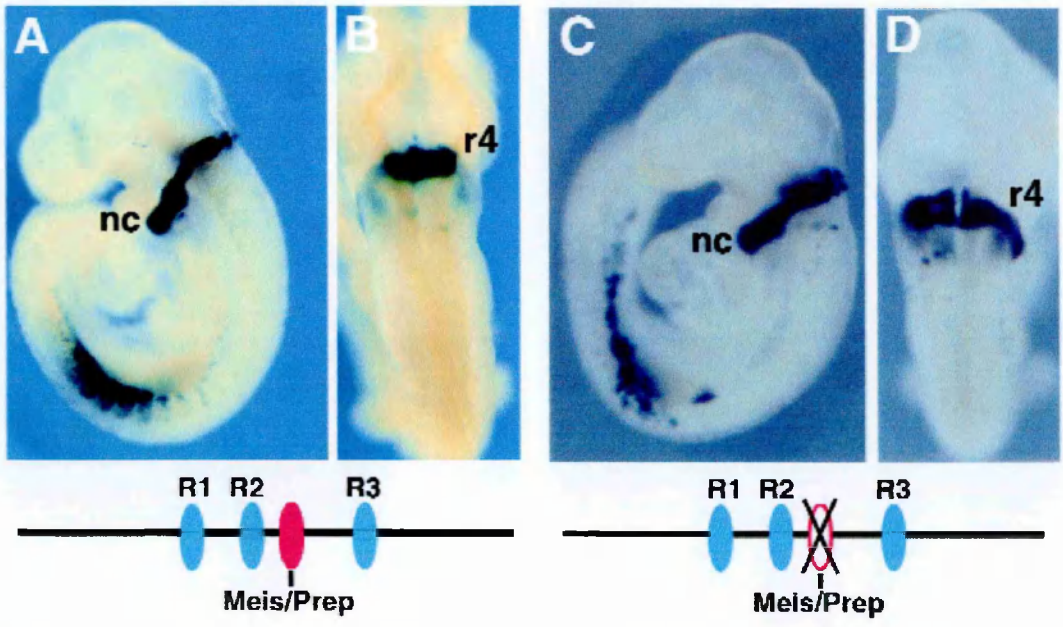


Figure 4.15

The PM site in the *Hoxb1* enhancer is not required for r4 expression. Lateral (A, C) and dorsal (B, D) views of a *Hoxb1* LacZ transgene in 9.5 d.p.c. embryos are shown. Blue: three bipartite PH sites (R1, R2 and R3); purple: PM site positionated beween R2 ans R3, An X indicates the presence of a mutation.
Strong r4-expression is mediated by the wild type version of a 331 bp StuI fragment of the *Hoxb1* enhancer (A, B). Neural crest (nc) can be seen to migrate out of r4 into the secon branchial arch. Identical expression is directed by a mutated enhancer in which mutations have been introduced into the PM site.

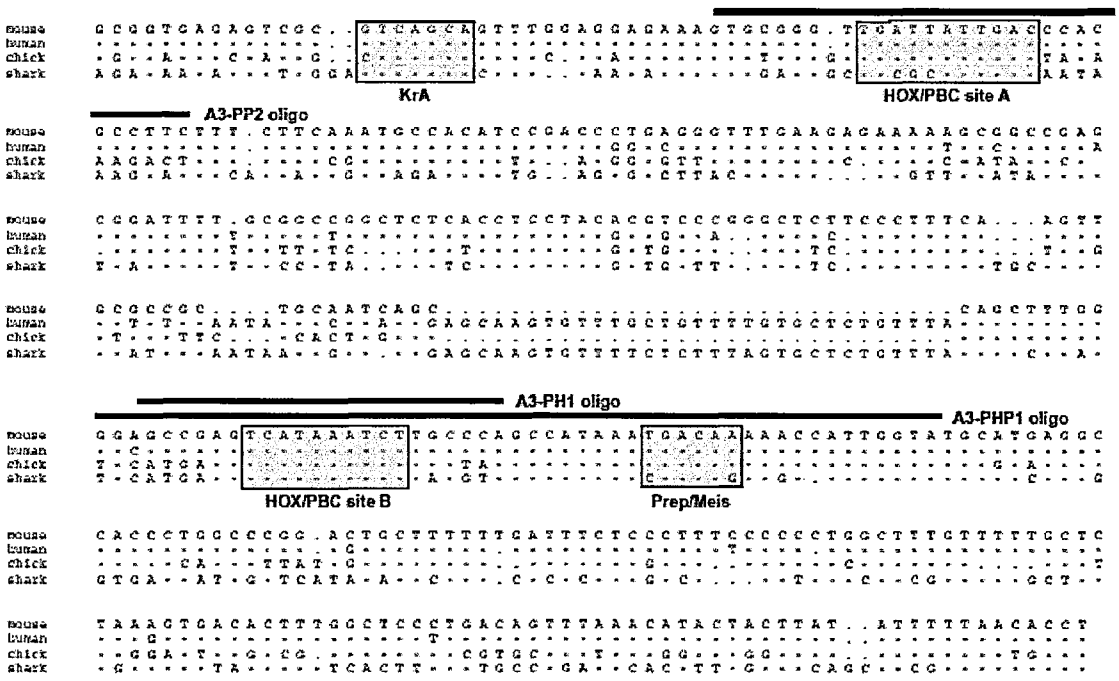


Figure 4. 16

Sequence alignment of a conserved region in the chick, mouse, human and horn shark *Hoxa3* enhancer. The Kreisler binding site is boxed and indicated by KrA. The putative Hox/PBC binding sites A and B, and the Prep1-Meis consensus motif, are boxed.
The black lines above the sequence indicate the double stranded oligonucleotide a3-PP2, a3-PH1, and a3-PHP1 used in binding and competition assays. Dashes represent identity and dots missing nucleotides (from Manzanares *et al.*, 1999)

The *r4-Hoxb1 enhancer* (331 bp StuI fragment Popperl *et al.*, 1995) is capable of directing the expression of the reporter gene in r4 and neural crest (Fig 4.15A-B). Previously I have demonstrated that PM-PH sequence of the *Hoxb1* enhancer bind the Prep1-Pbx-Hoxb1 ternary complex. However, on the contrary to the *Hoxb2* enhancer, no variation was observed in the expression pattern when the PM site in the *Hoxb1* enhancer was mutated (Fig 4.15 C and D). Thus, while the PM site in the *Hoxb2* enhancer is required for r4-expression, the PM site in the *Hoxb1* enhancer is not essential.

Similar data were obtained by Popperl *et al* (1995) showing that the multiple PH sites in *Hoxb1* enhancers are all required for r4 expression, while the mutation of a single PH site is not sufficient to change the reporter gene activity.

These results do not preclude a role for the PM site in the modulation of Hox expression. In fact, a synergism of the three PH sites might compensate for the mutation in the PM site in the *Hoxb1* enhancer while the single PH site in the context of the *Hoxb2* enhancer might not overcome the mutation. This aspect has been further explored in Chapter 5.

4.11 *Hoxa3* hindbrain *enhancer* contains functional Hox/Pbx binding sites.

Next I considered the cis-regulatory elements required for *Hoxa3* expression in r5/r6: once again in collaboration with Robb Krumlauf lab. It has previously been demonstrated that Kreisler regulates early sequential expression in r5/r6 by binding the *Hoxa3* enhancer (Manzanares *et al.*, 1999).

Manzanares showed that a relatively large block approximately 400 bp, located at 5-7 kb upstream of the Hoxa3 ATG in the *Hoxa3* mouse enhancer showed >70% of similarity with human chick and shark *Hoxa* enhancer sequences (Fig 4.16; Manzanares *et al.*, 2002). In addition to the well-known Kreisler (KrA) binding site Manzanares identified two putative Hox/PBC-(PH-A and-B) sites, related to consensus bipartite binding sites for Hox and Pbx proteins. In addition, a putative binding site for Pbx/Meinox heterodimers (PM site), located near a PH site like in *Hoxb1* and *Hoxb2* enhancers was also identified (Manzanares *et al* 1999 and 2002).

The bipartite PH A and B sites in the *Hoxa3* enhancer differ from the Hoxb1/Pbx target site previously identified. They contain TA or TT in the center of the core instead of GG, and they more closely resemble sites found in Hox 4-10 genes (Fig 4.16; Manzanares *et al.*, 2002).

Starting from the *Hoxa3* enhancer sequence, I designed a set of oligonucleotides to be used in EMSA: a3-PP2 oligonucleotide containing the PH-A binding site; a3-PH1 oligonucleotide containing the PH-B binding site and a3-PHP1 containing both PH-B and PM sites (Fig 4.16). In order to analyze the molecular interactions between Prep1, Pbx, Hox-3 proteins and the *Hoxa3* enhancer I used a competition assay to test if a3-PP2 and a3-PHP1 were able to interfere with the formation of complexes binding to b2-PM-PH in the presence of *in vitro* translated Prep1, Pbx1a and Hoxb1 proteins (Fig 4.17). The binding of Pbx1a-Hoxb1 and Prep1-Pbx1a dimers, and of the Prep1-Pbx1a-Hoxb1 ternary complex, was inhibited by the wild-type *Hoxa3* oligonucleotides (a3-PP2 and a3-PHP1) but not by the mutated oligonucleotides (MUT-A and MUT-B) containing the mutation in the Hox/PBC-A and in the Hox/PBC-B binding sites, respectively (Fig 4.17).

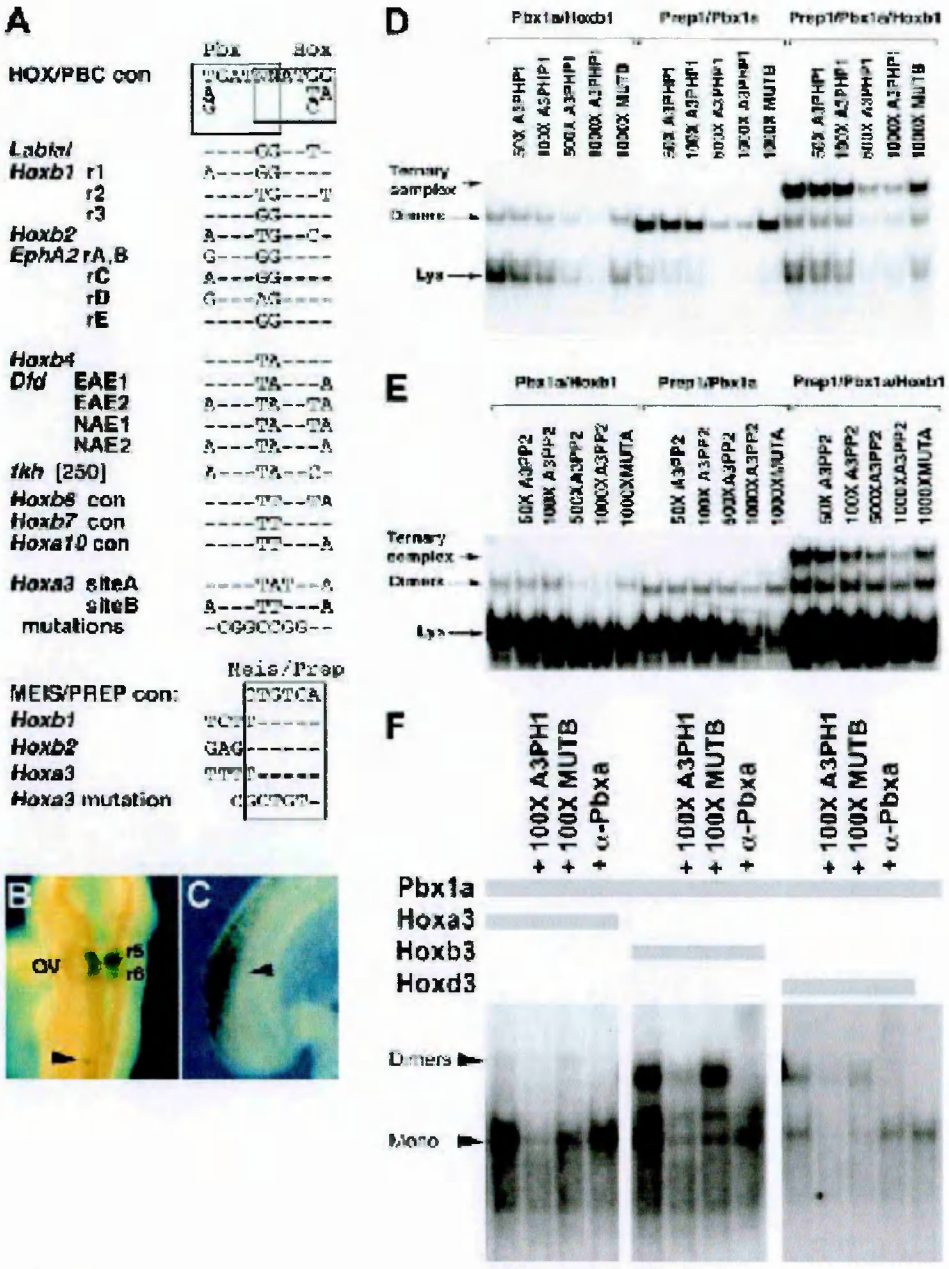


Figure 4. 17

Consensus bipartite Hox/PBC sites and their properties.

A)List of Hox/PBC (PH) and Meinox/Pbx (PM) binding sites. Mutation indicates the mutated oligonucleotide used in the binding assays.

B and C) Analysis of LacZ expression in 10.0 d.p.c. mouse embryos carrying a construct with five copies of Hox/PBC-A site linked to the reporter gene. In addition to r5/r6 expression the arrow marks posterior neural expression.

D and E) Competition assays which a labeled b2-PM-PH was mixed with Prep1, Pbx1a and Hoxb1 protein in the presence of 50-100- 500-1000 fold excess of cold a3-PP2 and a3-PM-PH oligonucleotide. MUT-A and MUT-B are mutated oligonucleotides (see panel A and methods). Arrows indicate dimeric and ternary complexes.

F) EMSA where labeled a3-PHP1 containing the PH site is mixed with Pbx1a, Hoxa3, Hoxb3 and Hoxd3 *in vitro* translated proteins. The binding of the complexes was competed with 100-fold excess of wild type (a3-PH1) and mutated (MUT-B) cold oligonucleotide.

The addition of anti-Pbx-a antibodies inhibits heterodimer formation (from Manzanares *et al.*, 1999).

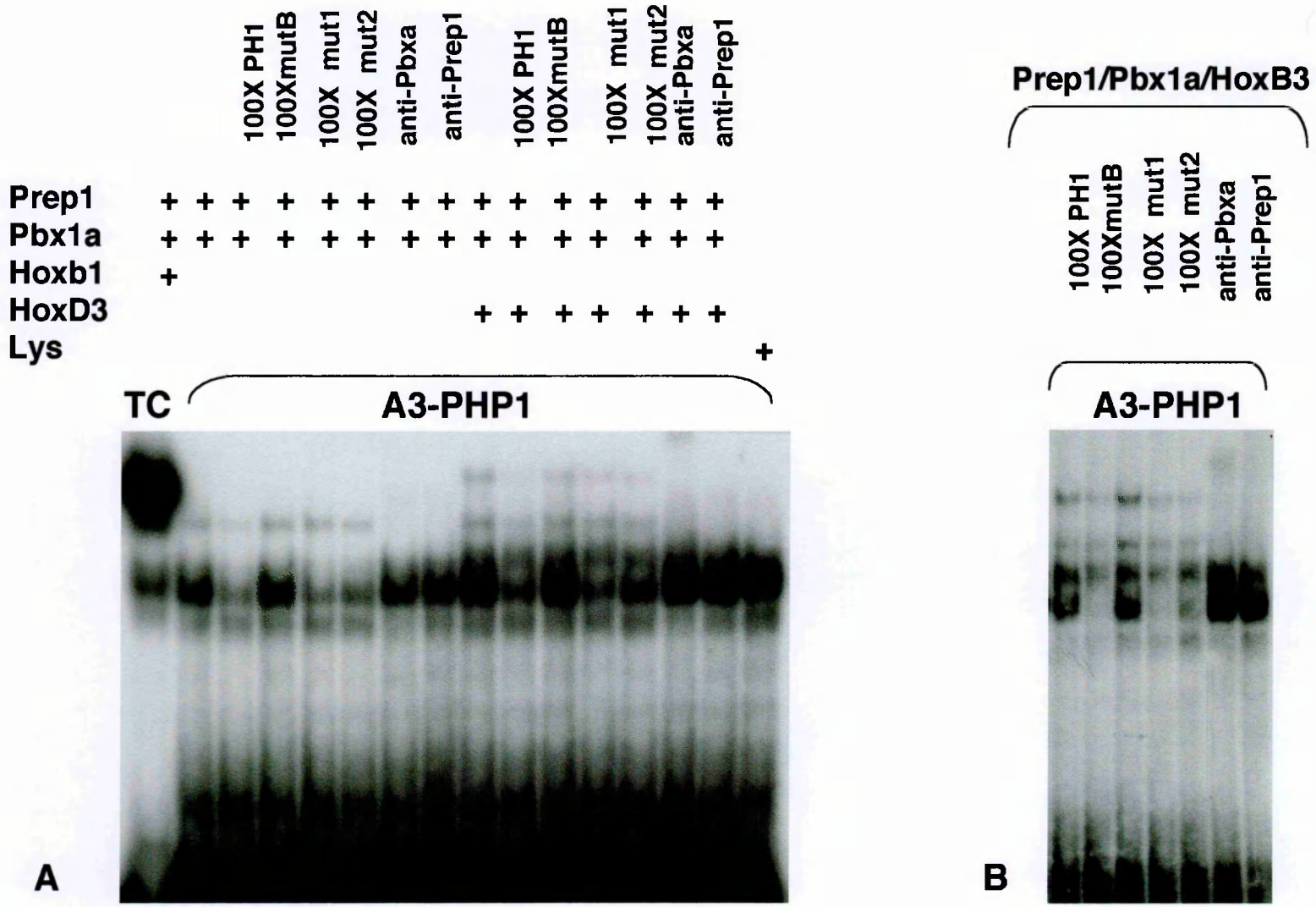


Figure 4.18
PM and PH sites of *Hoxa3* enhancer permit the assembly of Prep1-Pbx1a-Hoxd3 (panel A) and Prep1-Pbx1a-Hoxb3 (panel B) ternary complex The DNA-binding specificity of different complexes was tested by competition with wild type and mutated oligonucleotide and with antibodies anti-Prep1 and anti-Pbx-a. The symbol TC indicate the prep1-Pbx1a-Hoxb1 ternary complex formed on b2-PM-PH oligonucleotide

To test the direct binding of Hox group 3 proteins to Hoxa3 putative binding sites I performed EMSA using labeled *Hoxa3* oligonucleotides and *in vitro* translated Prep1, Pbx1a, Hoxa3, Hoxb3 and Hoxd3 proteins. In the presence of Pbx1a, Hoxa3, Hoxb3 and Hoxd3 proteins and labeled a3-PH1 oligonucleotide slower migrating bands appeared. These were formed by Pbx1a and group 3 Hox proteins (Fig 4.17F). Under these conditions Hoxb3-Pbx1a heterodimers are more efficient than Hox group 3 monomeric forms, with a relative efficiency of binding Hoxb3>Hoxd3>Hoxa3. The specificity of the complexes formed is demonstrated by the competition with wild-type and mutant *Hoxa3* oligonucleotide and using specific antibodies against Pbx-a proteins. Moreover, using a Hoxa3 oligonucleotide containing both PM and PH sites (a3-PHP1) I observed binding of Prep1-Pbx1a heterodimers and of the ternary complex Prep1-Pbx1a-Hoxb3 (Fig 4.18 B) and Prep1-Pbx1a-Hoxd3 (Fig 4.18 A). These *in vitro* results suggest a role Pbx-Hox group3 and for Prep1-Pbx complexes in the regulation of Hoxa3 hindbrain expression.
Moreover, they strongly support the cooperative role of PM and PH sites in the formation of the Prep1-Pbx-Hox ternary complex.

Manzanares *et al.*, (2001) have also demonstrated that Hox-PBC sites are required *in vivo* for the expression of Hoxa3 but only in the late phase, indicating Kreisler factor as the initial activator. Manzanares proposed that the initial activation of Hoxa3 expression by Kreisler was maintained by a Hox group 3 auto/cross-regulatory loop during late hindbrain development. This group has further demonstrated that single or combined mutation in the Hox/PBC binding sites reduced, but not totally abolished, the level of reporter gene expression in the late stage embryos, suggesting a synergic contribution of Hox/PBC binding sites (PH) to the late phase segmental Hoxa3 expression in r5/r6.

On the contrary, a mutation in the PM binding site taken together with a mutated KrA site had no effect on reporter-gene expression (Manzanares *et al.*, 2001). Therefore, *in vivo*, as is the case with *Hoxb1* enhancer, the contribution of the PM site for Hox expression is not required.
One reason for such variation might be that multiple Hox/Pbx sites in *Hoxa3*, like in *Hoxb1*, compensate for loss of the Pbx/Meinox motif, while the single Hox/Pbx site in the *Hoxb2* r4-enhancer cannot overcome this change.

Discussion chapter 4

During gastrulation and organogenesis in vertebrate embryos the Hox proteins play an important role in establishing basic body plan (Krumlauf 1994). Despite their central role in determining site-specific expression *in vivo, in vitro* they show low binding specificity. The DNA-binding ability and selectivity of these proteins depend upon the formation of specific heterodimeric complexes with members of the PBC family (Exd, Pbx).

Moreover, Hox proteins generally only show sequence conservation in the homeodomain and YPWM hexapepetide (Mann 1995). Despite these considerations they conserve similar function *in vivo* in different organisms, suggesting that these two domains are essential in determining protein function.

The homeodomain and the hexapeptide are important regions not only for the DNA binding but also to permit protein-protein interactions. Indeed, the interaction between Hox and PBC is mediated by the YPWM sequence of the Hox protein directly inserting into a hydrophobic pocket within the PBC homeodomain (Possner *et al.*, 1999; Piper *et al.*, 1999). PBC-Hox heterodimers specifically bind a bipartite consensus "TGATNNAT" identifies like a Pbx-Hox binding site or PH. In addition the Meinox proteins (Prep, Meis, Hth) interact with the PBC proteins and are able to modulate Hox activity.

Prep/Pbx/Hoxb1 proteins co-localize in r4 of the developing murine hindbrain.

As in the rest of the embryo, Hox proteins show a specifically restricted pattern of expression in the developing hindbrain. They are specifically expressed in portions of the hindbrain called rhombomeres. Rhombomere specific expression of many Hox proteins is dependent on *cis Hox* response elements.

In this thesis I demonstrate that Prep1/Pbx and Hoxb1 are present in r4 at 10.5 d.p.c (Fig 4.1). This observation provides initial evidence to support the involvement of Prep1 in regulating Hox expression. We previously demonstrated that *in vitro,* in transient transfection experiments the Pbx/Hox dependent transactivation of the autoregulatory element (ARE) of *Hoxb1* promoter is super activated by Prep1 transfection (Berthelsen *et al.*, 1998).

Formation of Prep1-Pbx-Hox ternary complexes on Hox target sites.

In this study I identify *cis* Prep/Meis PM elements near the bipartite PH (Pbx/Hox) binding sites in the *Hoxb1* and *Hoxb2* r4 enhancers and in the *Hoxa3* r5/r6 enhancer. I demonstrate that both PM and PH binding sites are required to modulate the formation of multimeric complexes containing Prep1-Pbx and Hox proteins.

Similar data were also reported for Hth-Exd-labial ternary complex that drives the expression of labial in the midgut endoderm cells, in Drosophila. Similarly the Meis1-Pbx-Hoxb1 ternary complex drives the expression of Hoxb2 in r4 in the mouse (Ryoo *et al.*, 1999; Jacobs *et al.*, 1999).

All these data suggest that the members of Meinox protein family play a crucial role in regulating the expression of *Hox genes* and that this mechanism is highly conserved.

PM-PH sequence like a general motif to modulate the *Hox genes* expression?

The presence of PM-PH elements in different *Hox* enhancers suggests that this might be a common model for many *Hox* response elements. Indeed, I demonstrate that the assembly of a Prep1/Pbx/Hox ternary complex occurs on all *Hox* enhancers considered, all of which contain and that had a PM-PH motif. Prep, Pbx and Hox proteins contain homeodomains and each of them significantly contributes to complex formation. Indeed, deletion of single homeodomains affects the binding activity. Thus, cooperative interaction throughout Meinox, PBC and Hox proteins and the combined PM-PH sites are essential to permit multimeric complex formation both *in vitro* and in cell culture.

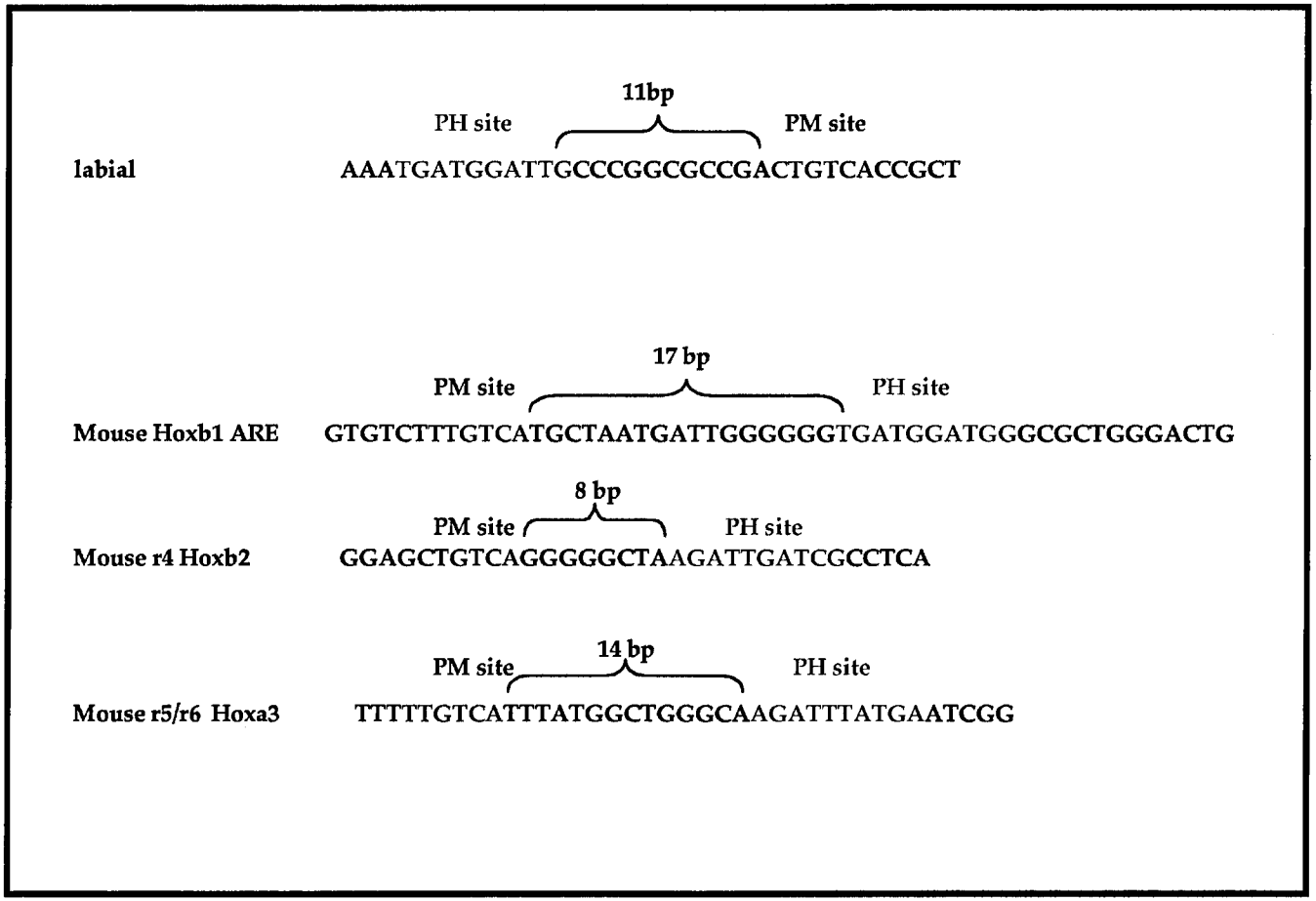


Figure 4. 19
Comparison of the contained PM-PH motif in *labial, Hoxb1, b1-ARE, r4-Hoxb2* and r5/r6 *Hoxa3* enhancers. The PM (Pbx/Meinox) binding site is indicated in blue, while the PH (Pbx/Hox) is shown in red. The distance between the two combined PM and PH sites are indicated in term of bp (base pair) on the top of the sequence.).

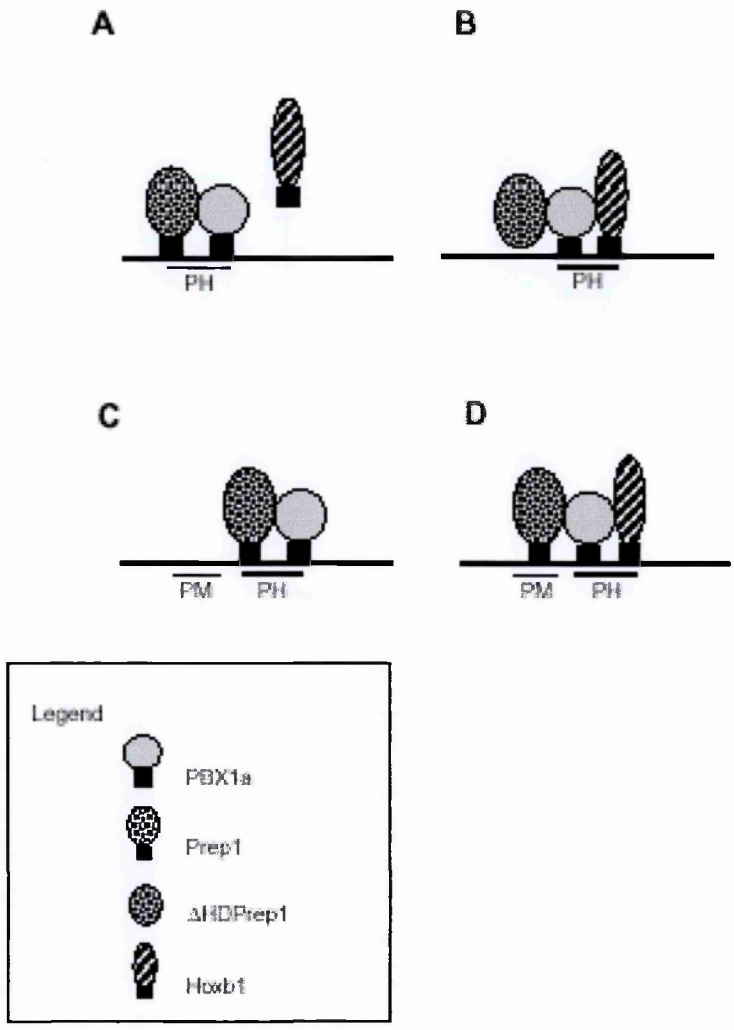


Figure 4.20

Ternary complex formation on PH vesus PM-PH sites in the *Hox* enhancer

A: Preformed Prep1-Pbx complex binds the PH site, since Hoxb1 does not bind Pbx in the absence of DNA. Prep1-Pbx may in part hide the formation of the Pbx-Hoxb1 complex on the PH site

B: Prep1 missing the homeodomain is still able to form a ternary complex.

C: In the case of the combined PM-PH motif a preformed Prep1-Pbx complex will prefentially bind to the PH site.

D: Hoxb1 may displace the DNA binding domain of Prep1 to the PM site and interacts with Pbx thus permitting ternary complex formation. Notice that in this case three homeodomains are bound to DNA.

It is interesting to note that arrangement, orientation and distance of the binding sites are different in all the enhancers studied (Fig 4.19). Firstly in labial, but not in *Hoxb1, Hoxb2* and *Hoxa3,* the PM site is downstream of the PH. Secondly, the distance between PM and PH is extremely variable, from 8 bp in *Hoxb2* to 17 bp in *Hoxb1.*

All these observations suggest that there may be significant flexibility in the assembly of the three proteins on these elements. This flexibility may be important for the recruitment of additional factors, which may be different in different enhancers. This topic will be considered in the detail in next chapter.
I demonstrate that it is a ternary complex that binds the PM-PH motif and not a double dimer, Prep1-Pbx and Pbx-Hoxb1. Thus, if a Prep1-Pbx-Hox ternary complex is binding the PM-PH motif, looping or DNA-bending could explain the ability of the ternary complex to bind combined PM-PH sites with different spacing. The bending could depend on the Prep-Pbx DNA-bending properties or may involve additional bending factors. For example, in the *Hoxb1* enhancer the linker region between PM and PH sites contains binding sites for the Oct1/Sox factors, which are known to have DNA-bending properties. However, we cannot exclude that Prep1 contains a flexible homeodomain able to extend and interact with distant sites therefore modifying the enhancer structures.

The induction of Hoxb1 in RA-treated P19 cells permits the formation of a ternary complex binding the combined PM-PH element.

Retinoid Acid (RA) is known to act as a morphogen molecule able to program hindbrain development (Conlon and Rossan, 1992; Marshall *et al.*, 1992). RA induces the expression of Hoxb1, *in vivo,* and the expression of Pbx proteins in cultured carcinoma P19 cells (Simeone *et al.*, 1990; Knoepler and Kamps, 1997). I show that Prep1 and low levels of Meis1a/b are present in the nuclei of un-induced P19 cells. After treatment with RA, the levels of Prep1 were not dramatically changed. On the contrary, Meis1a/b was strongly induced after 12 hours. Hoxb1 and Pbx expression were induced after RA treatment, as expected. Time course experiments show that Prep-Pbx is constantly present in the nuclear extract and constantly binds to the O1 oligonucleotide. On the contrary Prep-Pbx dimers bind to the PM-PH oligonucleotide only up until 6 hr after RA induction. After RA-induction the combined PM-PH oligonucleotide binds to a strong ternary complex. The ternary complex seems not to increase even in the presence of high amounts of Pbx and Meis. After 24 hr the levels of Hoxb1 decrease perhaps indicating that Hoxb1 is acting as a limiting factor. By dissociation rate experiments I also show that binding of Hoxb1 to the Prep1-Pbx complexes confers high stability to the ternary complex. Based on the results obtained with P19 cells, I can hypothesize that in the absence of Hoxb1 a pre-existing dimeric complex, Prep1-Pbx binds to the PH site (Fig 4.8). Recruitment of Hoxb1 by Pbx might displace the Prep1 homeodomain, inducing it to contact the adjacent PM site and permitting the formation of a ternary complex (Fig 4.20).

This model suggests that the Hox proteins may exert their function *in vivo* because of their ability to direct the assembly of sequence specific multimeric complexes containing highly conserved cofactors such as PBC and Meinox protein.

This model is in agreement with the hypothesis proposed by Mann *et al.*, (1998). These authors suggest that Hox, instead of TALE proteins, are the real cofactors and may be critical for the regulation of PBC containing complexes (Mann *et al.*, 1998).

Differences throughout the *Hox* enhancers

In vitro, the combined PM-PH sites are required to assemble the PBC-Meinox-Hox ternary complex on different *Hox* enhancers. The role of these binding sites *in vivo* seems to be different. I demonstrate an essential *in vivo* requirement for the PM site only in Hoxb2 expression in the developing hindbrain. This does not exclude a role for the PM site in the r4 expression of Hoxb1 or for the r5/r6 expression of Hoxa3 in the murine hindbrain. Indeed, comparing the structure of these three enhancers, we found that while *Hoxb1* and *Hoxa3* have respectively three and two PH binding sites, *Hoxb2* has only one.

It has been reported that the mutations introduced in single PH sites have no effect on reporter gene expression. These data seem to suggest the all the PH sites are cooperative binding sites that act in synergy. Indeed, all PH sites contributed to specifically restrict *Hox* gene expression. However, the presence of the PM site might still facilitate the ternary complex formation by cooperating with one of the other PH sites and therefore influencing the relative levels of expression important *in vivo* functions. In contrast, the *Hoxb2* enhancer has only a single PH-PM site and altering either component abolishes its activity. Thus, *Hox* target sites may vary in both number and arrangement of PH and PM sites and additional copies of either of these sites might serve in a redundant manner to ensure and reinforce levels of expression and spatial restriction or to modulate the ternary complex function. Moreover, the PM site contained in the *labial* enhancer shows a similar contribution in regulating the expression of labial in the Drosophila embryo (Ryoo *et al.*, 1999). It is been observed that point mutations in the PM site have no effect when introduced in the complete *labial* enhancer (*labial 550*), on the contrary the same mutations have a strong effect when introduced in the simple PM-PH motif. These data suggest that in the context of the entire *labial* enhancer there are other factors that stabilize the ternary complex, even in the absence of an optimal PM binding site. This property is probably applicable to the *Hoxb1* and *Hoxb3* enhancers, however this point requires further study in order to clarify the role of the ternary complex *in vivo*.

What is the role of the PM site and why is it necessary for Hoxb2 expression?

It is not clear why the PM site is required for Hoxb2 expression, however all the *in vitro* data strongly suggest that the PM site is essential to permit ternary complex formation. Prep1 does not have any apparent transactivation domain. In addition, the observation that the mutation in the *Hoxb2* PM site has a global effect on Hoxb2 expression suggests that the role of the PM site, and as a consequence the role of the ternary complex, might be that of mediator for the interaction with other factors binding the 2.1 kb sequence of the *Hoxb2* enhancer. Thus, the combined PM-PH sites and the bound ternary complex might create a bridge between different complexes. An extreme view is that those multimeric

complexes might act like adaptors able to create a connection with the basal transcription machinery complex.

DNA-binding specificity of PM site

It is interesting to note that while Prep1-Pbx dimers bind the PH site, they are unable to bind the PM site, which contains the canonical binding site for Prep-Pbx dimers (TGACAG). Prep1-Pbx heterodimers do not bind to the PM site in the context of a *Hox* enhancer. On the contrary, the same sequence in the Urokinase enhancer form Prep-Pbx heterodimeric complexes, suggesting that the bases flanking the PM sites, present in the *urokinase* enhancer, but missing in the Hox enhancers could modulate the DNA-binding properties. Furthermore, this observation suggests that the PM site could have different properties in different enhancer contexts. However Prep-Pbx binds also the PH site, containing the TGATNNAT motif.

The comparison of those different Prep1-Pbx binding sites led to the observation that Prep-Pbx complexes showed divergent DNA-binding specificity. So far the consensus PM must be redesigned in T/AGANNG (Fig 4.11). As shown in figure 4.11 the 3 bases (CTC) flanking the T/AGANNG consensus exert an essential role to convert a low affinity PM site to a high affinity site. Furthermore I have also demonstrated that the presence of high affinity PM site can alter the ternary complex formation. Indeed the b2-Z3 oligonucleotide containing both a high affinity PM and a PH sites form a weaker ternary complex (Fig 4.12 lane 10), suggesting that a high affinity PM site interferes with the formation of a ternary complex. This is probably due to the preferential binding of Prep-Pbx dimers that creates a steric hindrance hampering first the binding of Pbx-Hox dimers (lane 7).

In conclusion, these data suggest that depending on the sequence context two types of Prep1/Pbx binding sites can be distinguished:

1) High affinity sites, like the PM site in the *urokinase* enhancer the role of which is unknown.

2) Low affinity sites located near a Pbx/Hox binding site (PH), which are required for the ternary complex formation

The data presented in this chapter includes data obtained in collaboration with Robb Krumlaufs'group and Miguel Manzanares at MRC, National Institute for Medical Research, London, UK and are described in the following papers:

Ferretti, E., Marshall, H., Popperl, E., Maconochie, M., Krumlauf, R., and Blasi, F. (2000). "Segmental expression of Hoxb2 in r4 requires two separate sites that integrate cooperative interaction between Prep1, Pbx and Hox proteins" *Development* **127**: 155-166.

Manzanares, M., Bel-Vialar, S., Ariza-McNaughton, L., Ferretti, E., Marshall, H., Maconochie, M. M., Blasi, F. and Krumlauf, R .(2001). "Independent regulation of initiation and maintenance phases of Hoxa3 expression in the vertebrate hindbrain involve auto- and cross-regulatory mechanisms". *Development* **128**:3595-607."

CHAPTER 5
RESULTS AND DISCUSSION III

Does the arrangement of PH and PM binding sites affect ternary complex formation?

In both Hoxb1 and Hoxb2 enhancers, the bipartite PH (PH/Hoxb1) binding sites are involved in r4-restricted expression *in vivo* (Fig 5.1, Popperl; *et al* 1995). Both enhancers also contain one copy of the newly identified aPbx/Meinox (PM) site.

In vitro , an oligonucleotide containing both sites (PM+PH) is able to assemble a trimeric Prep1-Pbx1-Hoxb1 complex. Such a combined motif is present in both the *Hoxb1* and *Hoxb2* enhancers, although 17 and 8 bp separate the two sites, respectively. Both combined motifs are able to assemble trimeric complexes *in vitro*. *In vivo*, the mutation of the PM site in *Hoxb2* enhancer, but not that of *Hoxb1*, blocks LacZ expression in rhombomere 4 (see Chapter 4 and Ferretti *et al.*, 2000).

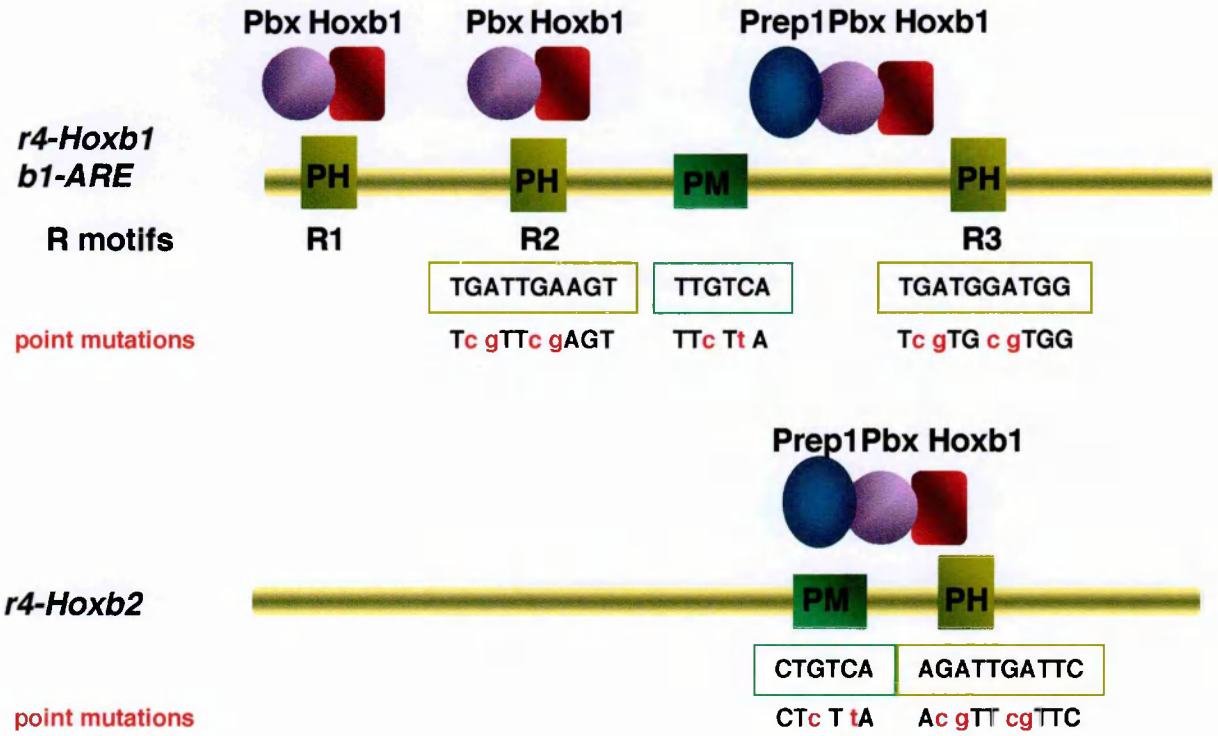
What are the differences between the *Hoxb1* and *Hoxb2* enhancers that can explain the *in vivo* results?

Despite the ability of the *Hoxb1* and *Hoxb2* enhancers to mediate similar patterns of r4-restricted expression *in vivo*, they differ in the organization and number of their bipartite PH (Pbx/Hoxb1) binding sites (Fig 5.1; Popperl, *et al.*, 1995). The *Hoxb2* enhancer has one PH site (Maconochie *et al.*, 1996), while the *Hoxb1* enhancer has three PH sites called R motifs (R1, R2 and R3) which are located within a highly conserved 331 bp region (*b1-ARE*).

A synergism between the three PH sites might compensate for the mutation in the PM site, while the single PH site in the context of *Hoxb2* enhancer may not overcome this change. *In vivo*, the synergisms between the three PH sites might allow the ternary complex formation in the absence of a PM site. To support this theory it has been shown that when one of the PH sites is mutated the activity of the enhancer is not totally lost, but only reduced. The major effect was seen with the introduction of point mutations in R3, while mutations in R2 and R1 had only a modest effect (Popperl *et al.*, 1995). Thus, the combined PM-PH site and the presence of multiple PH binding sites might serve in a redundant manner to ensure or reinforce levels of expression and/or spatial restrictions. Alternatively, the presence of multiple PH sites in the *Hoxb1* enhancer may directly affect the ternary complex by preventing its formation. An *in vitro* molecular approach has been used to investigate these opposite possibilities.

5.1 Can the R2 site substitute for the PM site in the formation of the ternary complex?

In order to answer this question, I performed EMSA using a set of wild-type and mutated oligonucleotides, derived from the *Hoxb1* enhancer and *in vitro* translated proteins (Fig 5.2). As a positive control I used the combined PM-PH



Hoxb1/Pbx 1 consensus	TGATNNATNN
Hoxb1-R 1	AGATGGATGG
Hoxb1-R 2	TGATTGAAGT
Hoxb1-R 3	TGATGGATGG
Hoxb2-PH	AGATTGATTC

Figure 5.1
Schematic representation of *r4-Hoxb1* and *r4-Hoxb2* enhancers. The sequences of the PH binding sites are boxed in green. Red bases indicate the point mutations introduced into the binding sites.
Alignment of *r4-Hoxb1* (repeats 1, 2 and 3)and *r4-Hoxb2* PH binding sites and comparison with the Pbx-Hoxb1 consensus are indicated. The variable bases are boxed in red.

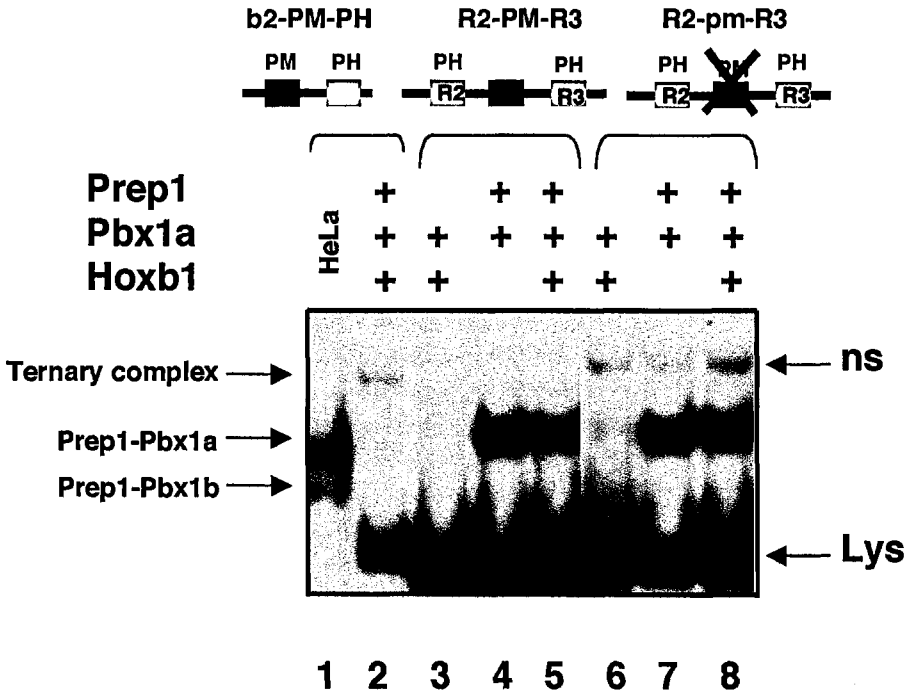


Figure 5.2

The addition of the R2 site to the PM-R3 oligonucleotide inhibits ternary complex formation. EMSA was performed using a combination of *in vitro* translated Prep1, Pbx1a and Hoxb1 proteins, as indicated. The control b2-PM-PH oligonucleotide form a ternary Prep1-Pbx-Hoxb1 complex, while both the R2-PM-R3 and the R2-pm-R3 oligonucleotides are unable to form the ternary complex.

The labeled DNA targets are shown on the top. The sequence of oligonucleotides and the nature of mutations are shown in Fig 5.1. Reticulocyte lysate contains non-specific factors forming a slower migrating complex marked ns and a fast migrating one marked Lys.

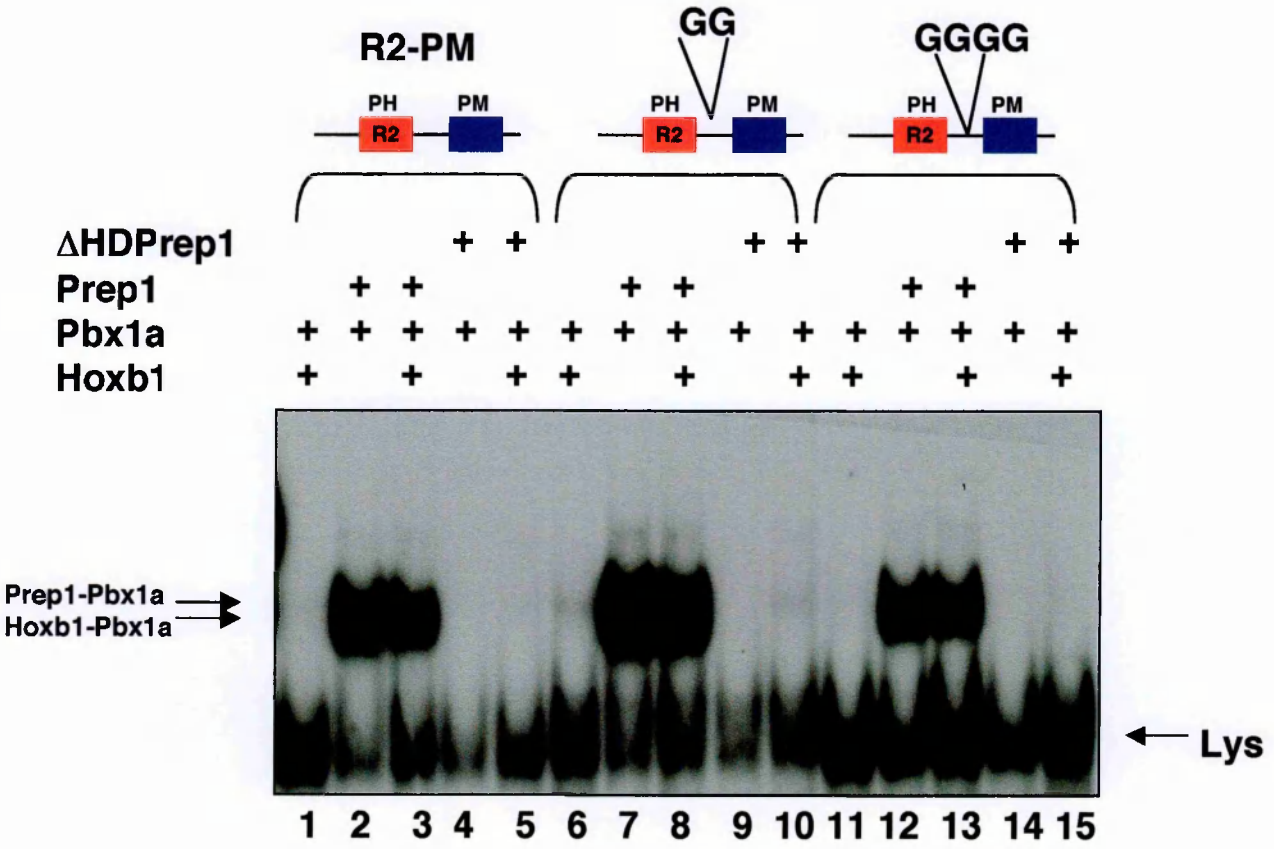


Figure 5.3
R2 is unable to cooperate with the PM to form a ternary complex. The PM-R2 oligonucleotide binds Prep1-Pbx1a dimers with high affinity. Insertion of 2 or 4 bp has no influence on the binding properties. EMSAwas performed using combinations of *in vitro* translated Prep1, Pbx1a, Hoxb1 and ΔHDPrep1 proteins, as indicated. The labeled DNA targets are shown on the top.

site from the *Hoxb2* enhancer. This oligonucleotide (b2-PM-PH) is able to bind Prep1-Pbx dimers if incubated with HeLa nuclear extracts (Fig 5.2 lane 1). In the presence of Prep1, Pbx1a and Hoxb1 *in vitro* translated proteins, b2-PM-PH forms a ternary complex (Fig 5.2 lane 2; Fig 4.3A.; Ferretti *et al.*, 2000). Subsequently the binding activity of b2-PM-PH was compared with the binding of two oligonucleotides derived from *Hoxb1* enhancer: the wild type R2-PM-R3 oligonucleotide in which two PH sites (R2 and R3) flank the PM site, and the mutated R2-pm-R3 oligonucleotide carrying a mutation in the PM site (Fig 5.2). Both wild type and mutated *Hoxb1* oligonucleotides strongly bind the Prep1-Pbx1a heterodimer (lanes 4 and 7) and more weakly the Pbx1a-Hoxb1 heterodimer (lanes 3 and 6). Surprisingly, no ternary complexes are detected when Prep1, Pbx1a and Hoxb1 proteins are present at the same time, suggesting that the presence of the R2 site interferes with the formation of the ternary complex (lanes 5 and 8). Moreover, the mutation in the PM site increased the binding activity of an endogenous, unidentified, more slowly migrating factor, present in the reticulocyte lysate and marked non-specific (ns) in the figure 5.2 lanes 6, 7 and 8.

The binding properties of the R2 and R3 sites were next compared. EMSAs were performed using an oligonucleotide containing the R2 and PM sites, but not R3 (Fig 5.3). I found that the R2-PM oligonucleotide weakly binds Pbx1a-Hoxb1 dimers and strongly associates with the Prep1-Pbx1a heterodimers (lanes 1 and 2, respectively). However, when all three proteins, Prep1, Pbx1a and Hoxb1 were present the R2-PM oligonucleotide continued to bind the Prep1-Pbx1a dimers, but no ternary complexes were formed (lane 3).

Since we know that R2-PM sites are separated by only 6 bp, whereas PM-R3 in *Hoxb1* and PH-PM in *Hoxb2* are separated by 17 and 8 bp respectively, I decided to investigate the relevance of the physical distance between R2 and PM sites on ternary complex formation (Fig 5.3). I found that, in the presence of all three proteins (Prep1, Pbx1a, and Hoxb1), the R2 and PM sites were not able to cooperate and form the ternary complex even when 2 or 4 additional bases were introduced between the two sites (lanes 8 and 13). In addition, however both mutated oligonucleotides are still able to bind the Prep1/Pbx dimers with high affinity (lanes 7 and 12), since the R3 site forms a ternary complex in the presence of ΔHDPrep1, tested if R2 might have the same property: in the presence of ΔHDPrep1, Pbx1a and Hoxb1 were unable to form a ternary complex and bind the R2-PM binding site (Fig 5.3 lanes 5, 10 and 15).

In conclusion, the R2 site may bind Prep1-Pbx dimers, but is unable to cooperate with the PM site to form a ternary complex. It may also interfere with the formation of the TC on the PM-R3 combined site.
The binding activity and cooperative properties of R2 and R3 appear to be different.

5.2 The R2 site inhibits the formation of ternary complex on the b1-ARE element.

I investigated the effect of the R2 site on the ternary complex formation by the otherwise active PM-R3 sequence. To this end I used a R2-PM-R3 oligonucleotide mutated in the R2 site (r2-PM-R3). I found that the oligonucleotide not only binds the heterodimers Pbx1a-Hoxb1 and Prep1-Pbx1a

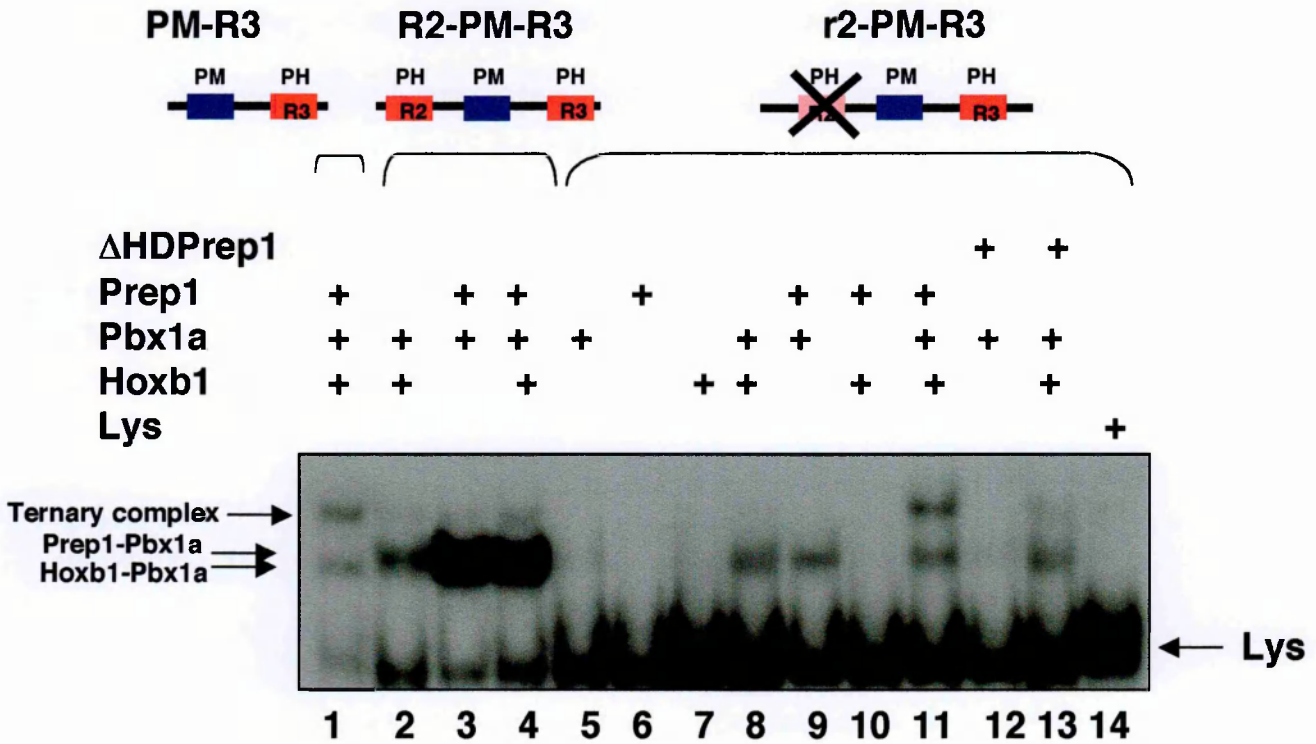


Figure 5.4
Mutations in R2 restore the ability of PM and R3 to cooperate in the formation of the ternary complex. *In vitro* translated Prep1, Pbx1a and Hoxb1 form a ternary complex on the mutated r2-PM-R3 oligonucleotide (lane 11), but not on wild type R2-PM-R3 (lane 4), which is only able to bind Prep1-Pbx dimers (lane 3). The b1-PM-R3 oligonucleotide was used as positive control for ternary complex formation (lane 1).

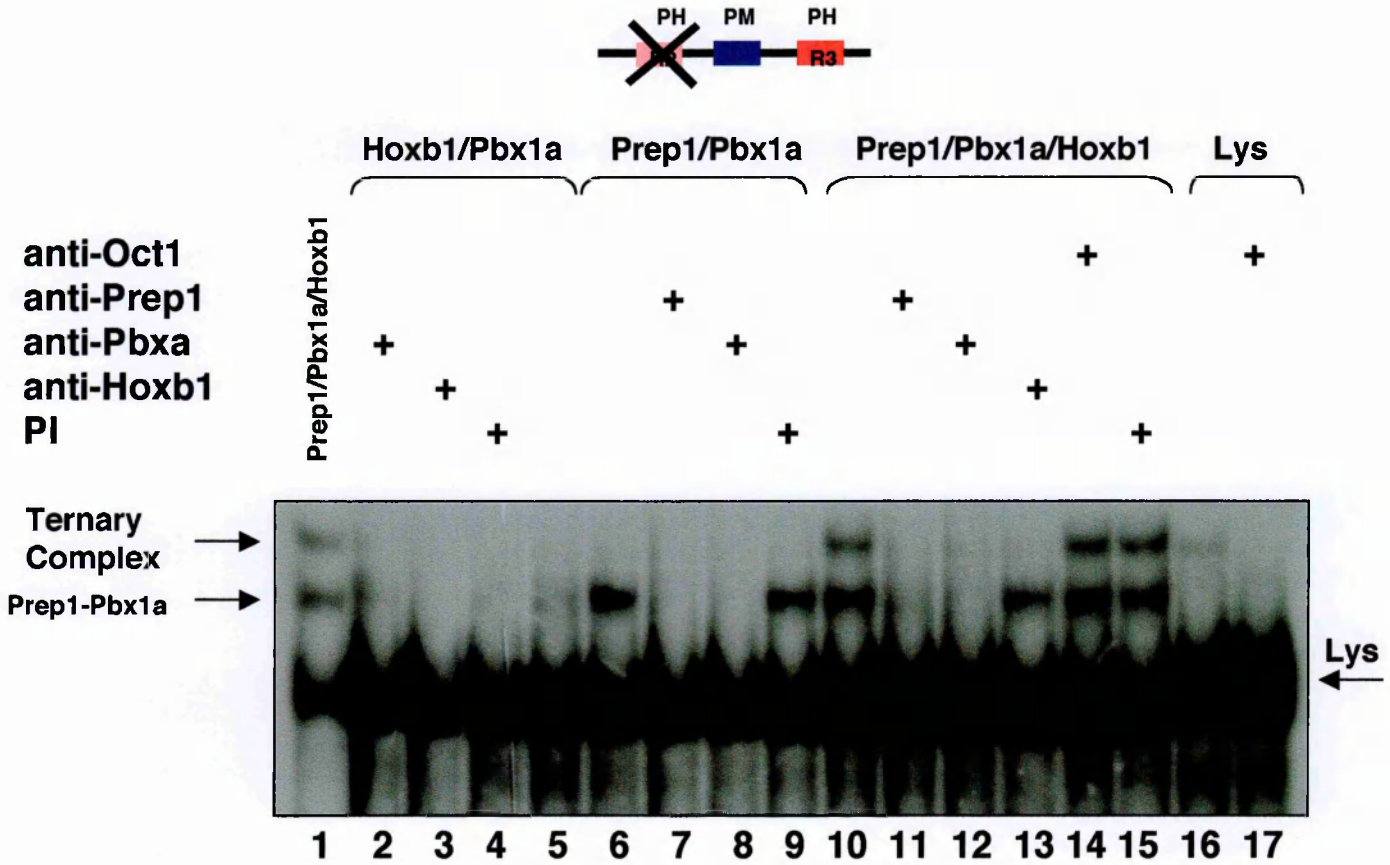


Figure 5.5
The ternary complex Prep1-Pbx- Hoxb1 and the Oct1 transcription factor bind the mutated r2-PM-R3 oligonucleotide but produce two distinct co-migrating bands. Addition of anti-Pbx, anti-Prep1 and anti-Hoxb1 antibodies inhibits the DNA-binding activity of the ternary complex. Anti-Oct1 antibody inhibits the endogenous reticulocyte band. The antibodies and oligonucleotide used in EMSA are indicated on the top.

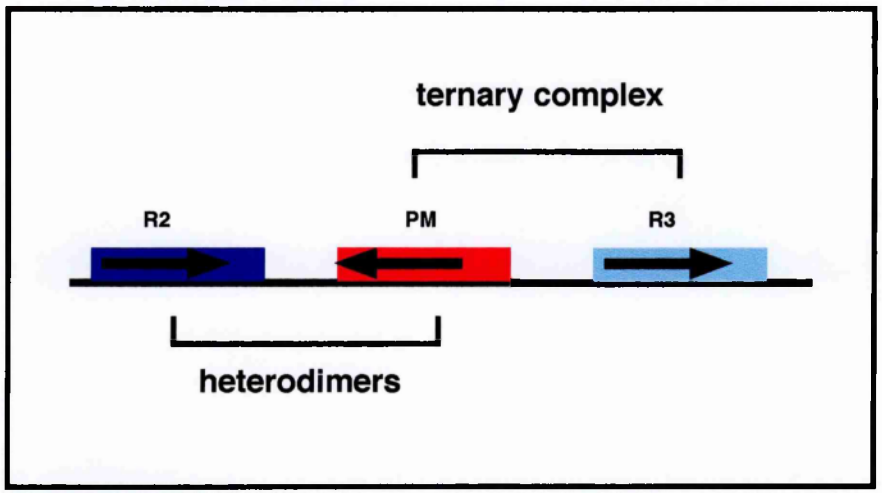


Figure 5.6
Schematic representation of the molecular interactions occurring on the PH-PM sites in the *Hoxb1* enhancer.

(Fig 5.4, lanes 8 and 9), but also a multimeric complex that co-migrates with the ternary complex formed on the PM–R3 oligonucleotide. These results confirm the inhibitory effect of R2. When R2 is mutated, R3 is still able to cooperate with PM and allow the formation of the ternary complex, suggesting that in the context of the *Hoxb1* enhancer there is a cross-talk between R2 and the cooperating PM-R3 element (Fig 5.1, lane 11). Indeed, an active R2 site inhibits the cooperation between R3 and PM.

In order to demonstrate the specificity of the bands observed in figure 5.4, I repeated the EMSA using specific antibodies (Fig 5.5). Anti-Pbx-a and anti-Hoxb1 antibodies inhibited Pbx1a-Hoxb1 heterodimer formation (lanes 3 and 4); anti Pbx-a and anti Prep1 antibodies inhibited Prep1/Pbx1a heterodimer formation (lanes 7 and 8), while anti-Prep1, -Pbx1a or -Hoxb1 antibodies prevented the formation of the ternary complexes (lanes 11, 12 and 13).

Previous studies have shown that other factors, besides Pbx, Hoxb1 and Prep1, can bind the region between R2 and R3 (Di Rocco *et al.*, 2001). One of these factors is Oct1, which binds the octamer-like sequence ATGCTAAT located next to R3 (Di Rocco *et al.*, 2002). Since I observed that Oct1 is very abundant in the reticulocyte lysate I wanted to test if the endogenous binding activity of the reticulocyte lysate, which interferes with ternary complex identification, corresponded to Oct1. Indeed, the slow migrating band in the reticulocyte lysate was completely inhibited by an anti-Oct1 antibody (Fig 5.5, lane 16). However, the addition of anti-Oct1 antibody did not affect the formation of the ternary complex (lane 14). I can therefore conclude that Oct1 and the Prep1-Pbx1a-Hoxb1 ternary complex produce two distinct, co-migrating bands. This becomes relevant for the subsequent experiments (see paragraph 5.4).

5.3 Influence of the spacing and the orientation of the R2, PM and R3 binding sites on ternary complex formation.

The experiments described above allowed me to conclude that the R2 binding site in the *Hoxb1* enhancer may prevent the formation of ternary complex *in vitro*.

In order to understand how R2 prevents the formation of the ternary complex and how the PH sites are able to determine which complex will form, I performed mutagenesis analysis of the *Hoxb1* enhancer.

R2-PM was shown to be a high affinity binding site for Prep1-Pbx1a heterodimers, and that its presence influenced the cooperation between PM and R3 in the ternary complex formation (Fig 5.6). I further analyzed the influence of the spacing between PM and R3 sites and the relative orientation of R2, PM and R3 sites on ternary complex formation *in vitro* . First I tested if R2 is able to mediate ternary complex formation if it is located downstream of PM. This was achieved using the PM-R2 oligonucleotide termed Flop1. I then investigated if R3 behaves like R2 when moved upstream of PM, using the R3-PM oligonucleotide termed Flop2. The wild type PM-R3 and R2-PM oligonucleotides were used as controls (Fig 5.7). The EMSA revealed that the R2 site is not able to cooperate with the PM site to form a ternary complex even if it is located in the same orientation and at the same distance as R3 (Fig 5.7, compare lanes 3 and 9). In fact, no slower migrating bands were detected in the presence of Prep1, Pbx1a and Hoxb1 proteins. On the other hand, moving R3 upstream of PM did not

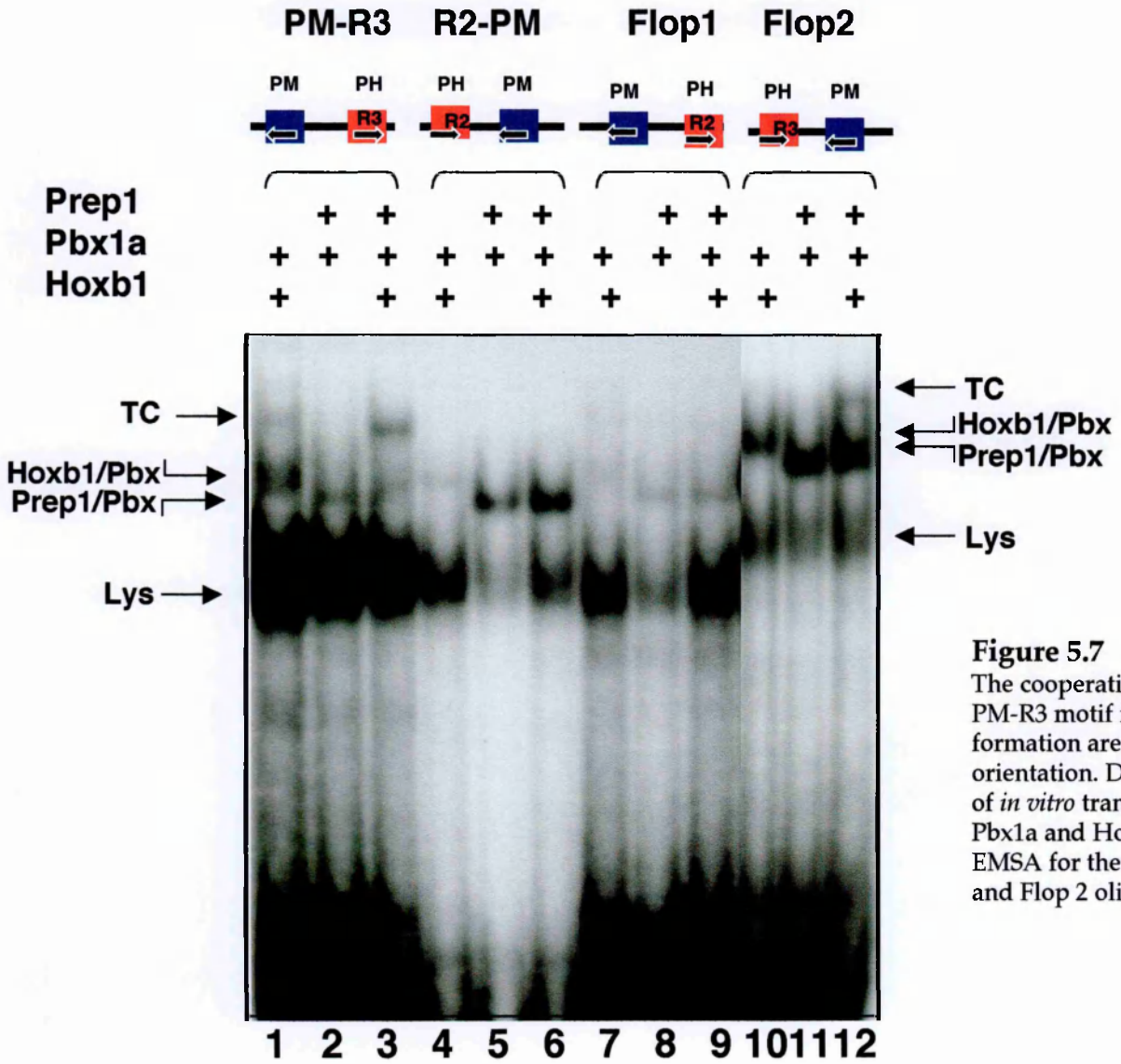


Figure 5.7
The cooperative properties of the PM-R3 motif in ternary complex formation are independent of their orientation. Different combinations of *in vitro* translated proteins Prep1, Pbx1a and Hoxb1 were tested by EMSA for their ability to bind Flop1 and Flop 2 oligonucletides .

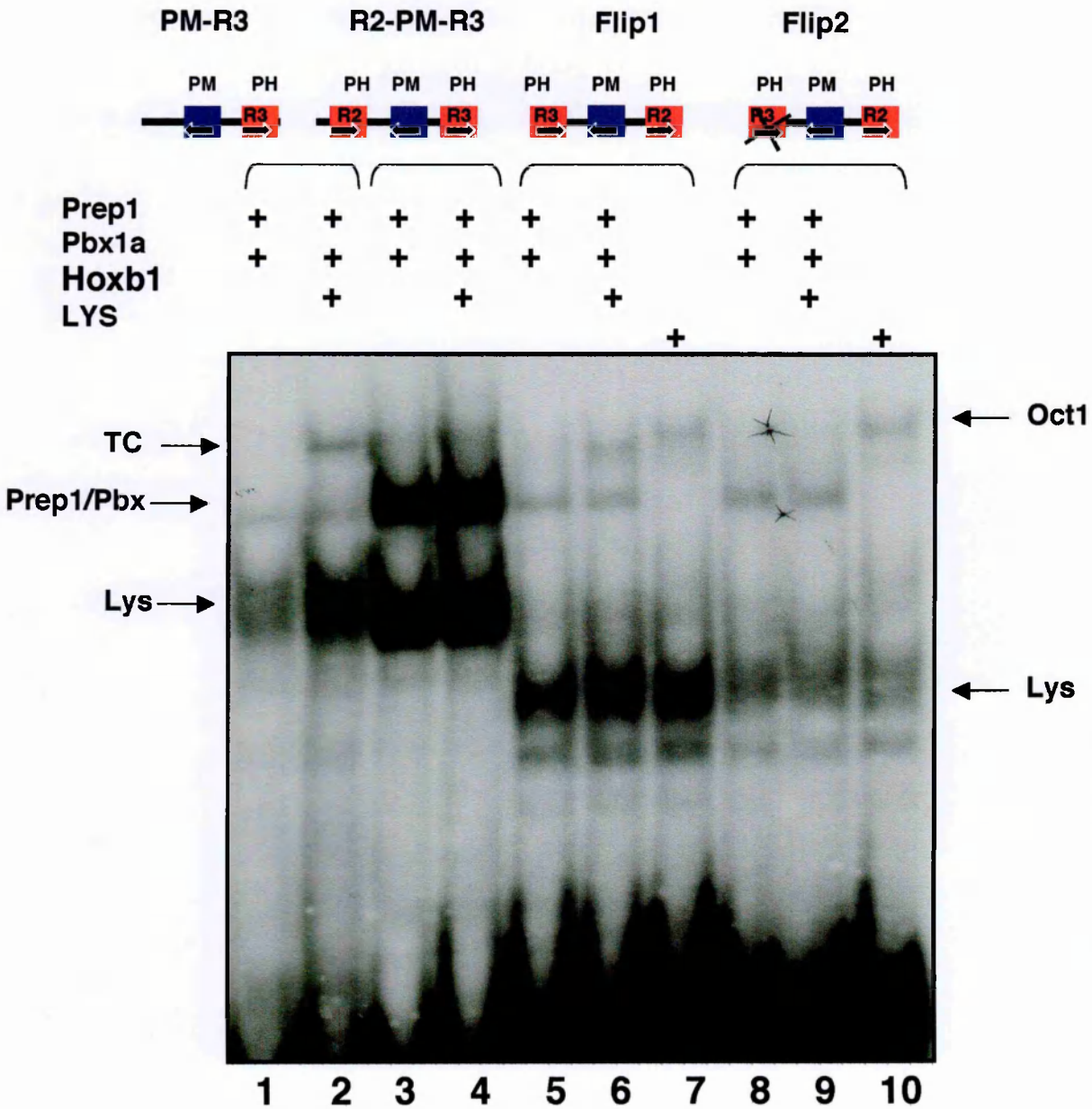


Figure 5.8
Binding properties of *in vitro* translated Prep1, Pbx1a and Hoxb1 proteins on Flip1 and Flip2 oligonucleotides. The modifications and oligonucleotides used are indicated at the top . The activities present in reticulocyte lysate are indicated by Lys for unspecific factor and with Oct1 for the characterized binding factor.

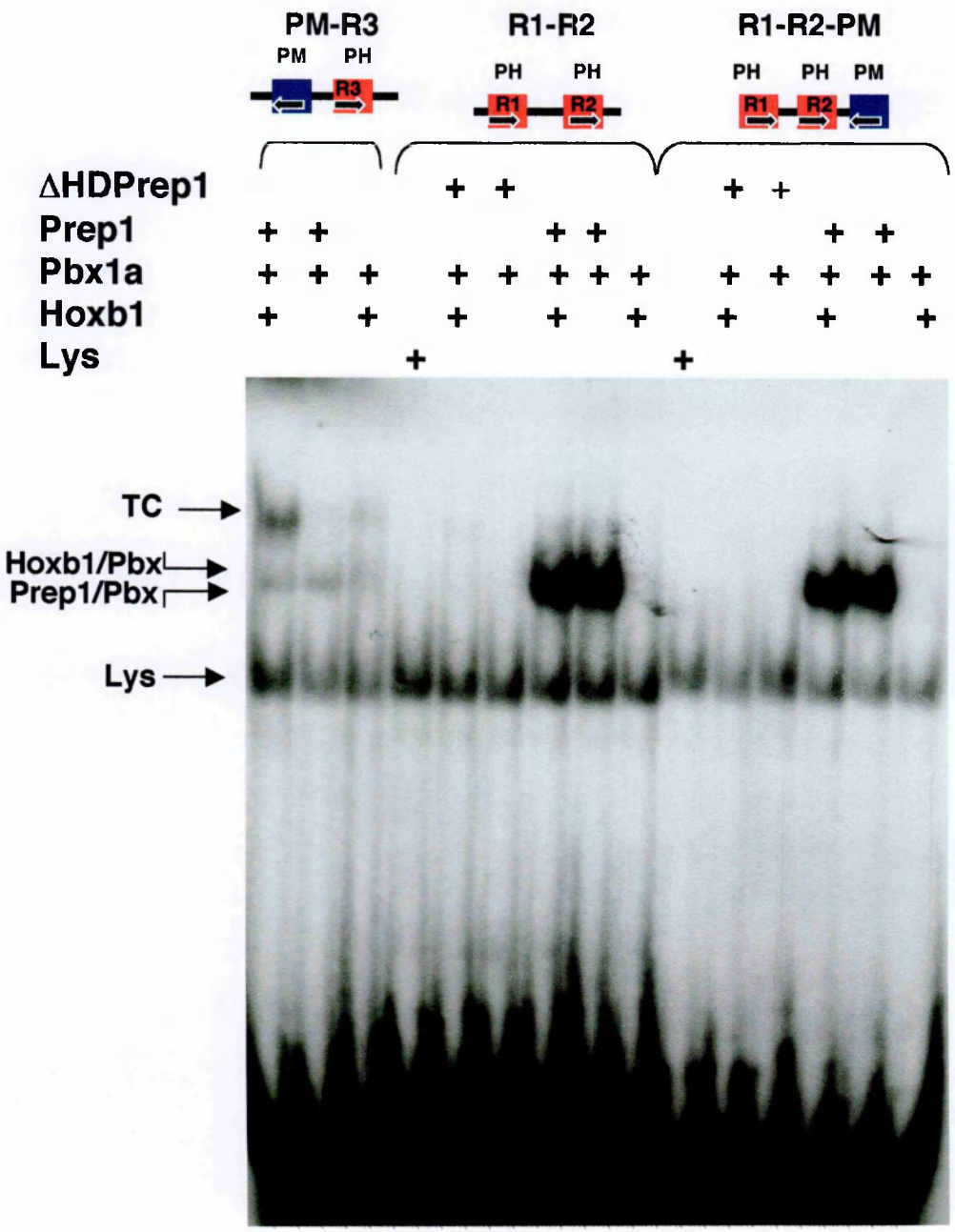


Figure 5.9
Both R1-R2 and R1-R2-PM oligonucleotides bind with high affinity Prep1-Pbx dimers, but are both unable to cooperate with PM to form a ternary complex . EMSA was performed using combinations of *in vitro* translated Prep1, Pbx1a, Hoxb1 and ΔHDPrep1 proteins, as indicated. The labeled DNA targets are shown on the top.
The activities present in reticulocyte lysate are indicated by Lys.

affect the formation of the slower migrating band (lane 12) suggesting that changing R3 space-orientation has no influence on the cooperation between the R3 and PM sites. I also investigated if the cooperative properties of PM-R3 were still inhibited by R2 when R2 and R3 are flipped. The EMSA in figure 5.8 demonstrates that flipping of R2 and R3 position (Flip 1 oligonucleotide) results in lower binding affinity for Prep1-Pbx1a heterodimers (compare lanes 3 and 5) and restores the formation of the ternary complex (compare lanes 2, 4 and 6) suggesting that changing the relative position of R2 and R3 reduces the inhibitory properties of R2. Moreover, Fig 5.8 shows that the presence of wild-type R3 is essential for ternary complex formation as mutations in R3 (Flip2) prevent its formation (lane 9). Like the wild type oligonucleotides, Flip1 and Flip2 bind the Oct1 ubiquitous factor present in the reticulocytes lysate (lanes 7 and 10).

These experiments show that the spacing and orientation between PM and R3 sites have no influence on ternary complex formation *in vitro*. In contrast, the cooperation between the PM and R3 sites, in the context of the *Hoxb1* enhancer, is inhibited by the presence of R2. Hence, R2 may influence dynamic interaction and total activity of the *b1-ARE*.

5.4 R1 and R2 binding motifs bind Prep1-Pbx dimers but cannot interact with PM to form a ternary complex.

I also analyzed the molecular interactions between Prep1/Pbx/Hoxb1 and two other PH sites of the *r4-Hoxb1* enhancer, as this may be important in modulating the cumulative enhancer activity. I performed EMSA using an oligonucleotide containing the R1 and R2 sites, or an oligonucleotide containing R1, R2 and PM, but missing the R3 site (Fig 5.9).

I found that the R1-R2 oligonucleotide weakly bound Pbx1a-Hoxb1 dimers but it bound with high affinity to the Prep1-Pbx1a heterodimers. Moreover, when all three proteins (Prep1, Pbx1a and Hoxb1) were present, the R1-R2 oligonucleotide continued to bind the Prep1-Pbx1a dimers, but no ternary complex was formed (Fig 5.9). R1-R2-PM oligonucleotide showed the same DNA-binding activity.

In conclusion, both R1 and R2 sites bind Prep1-Pbx dimers but are unable to cooperate with the PM site to form a ternary complex. Therefore the binding and cooperative properties of R1 and R2 differ from those of R3. For all these reasons we should rename R1 and R2 as PM instead PH binding sites

5.5 Nuclear extracts of P19 cells contain constitutive Oct1 and an inducible Prep/Pbx/Hoxb1 ternary complex.

To characterize the binding properties of the oligonucleotides described above in a cellular context I prepared nuclear extracts from P19 cells maintained under basal conditions or after induction with retinoic acid (RA).

Previous experiments revealed that the b2-PM-PH constitutively binds the Prep1-Pbx heterodimer present in nuclear extracts of untreated cells and the Prep1-Pbx-Hoxb1 ternary complex formed after induction with retinoic acid (RA) (Fig 4.8A and Ferretti *et al.*, 2000).

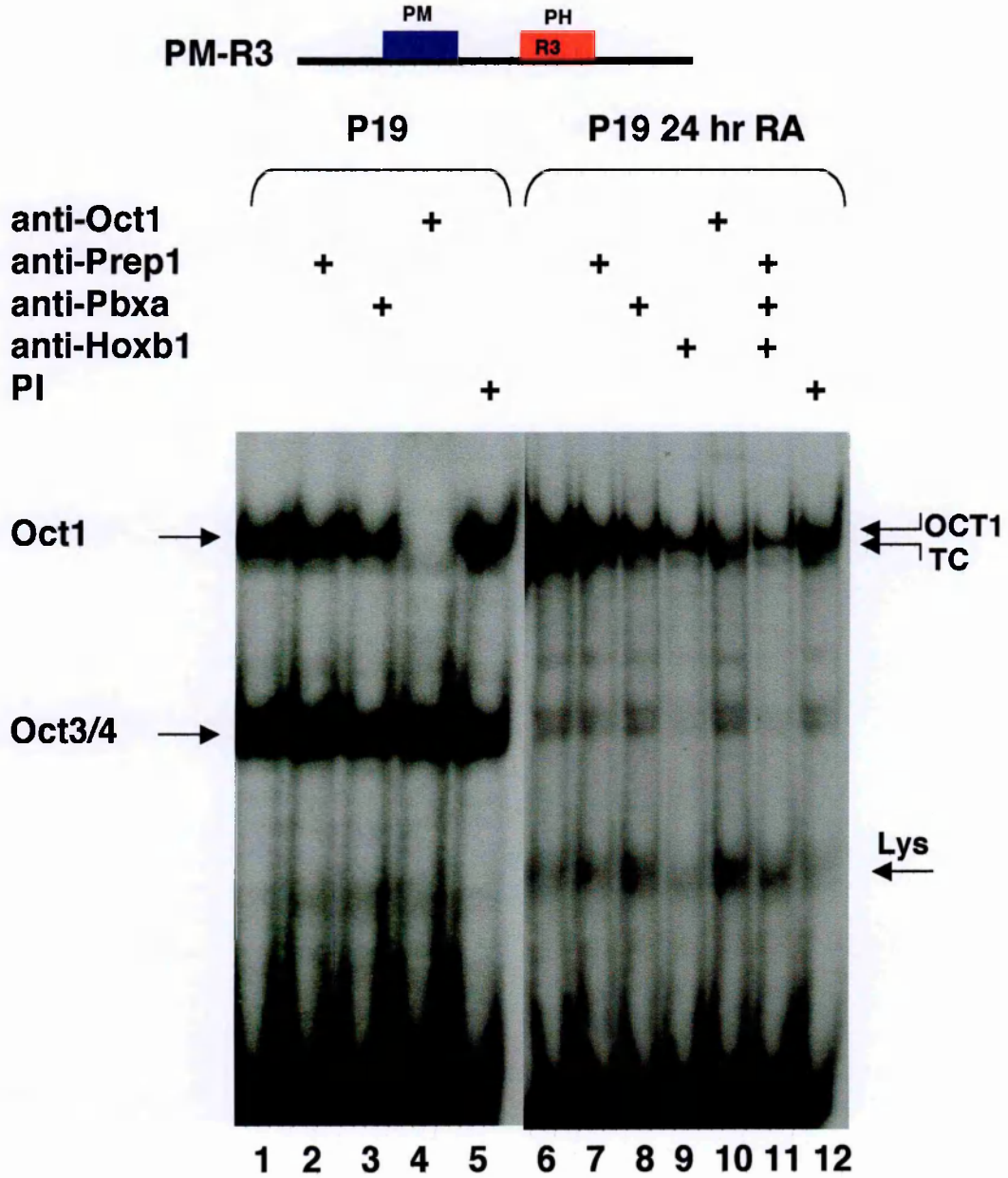


Figure 5.10
Ternary complex formation in untreated and RA-induced P19 nuclear extracts using the PM-R3 oligonucleotide. Notice that PM-R3 (PM+PH) is missing the R2 binding site. The binding specificity was tested using specific antibodies as indicated on the top. Nuclear extracts from P19 cells contain constitutive Oct1 (lanes 4 and 11) and an inducible Prep1-Pbx-Hoxb1 ternary complex (TC). They also contain a down-regulated Oct3/4 binding activity (compare lanes 1 and 6).

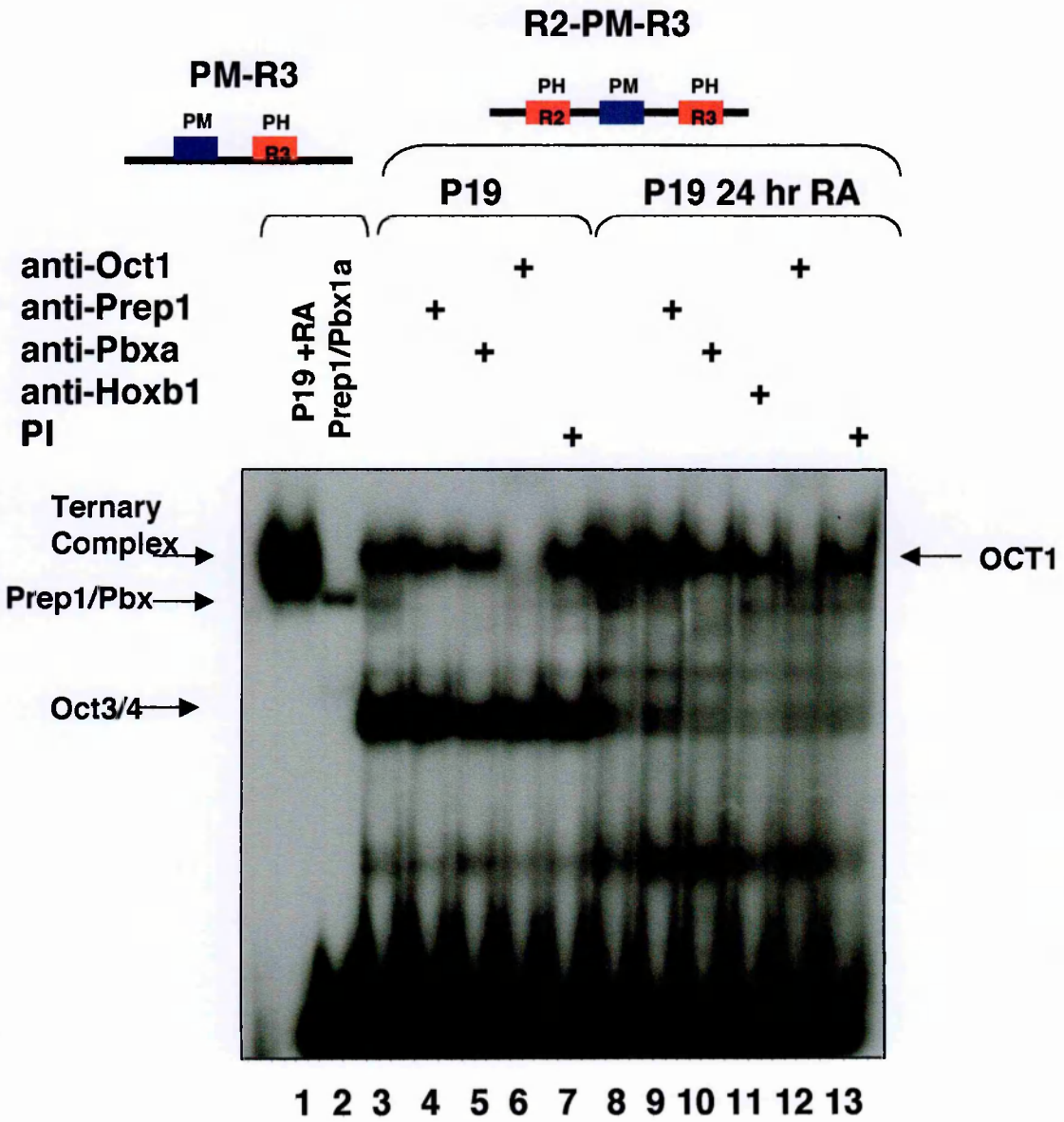


Figure 5.11
The presence of R2 facilitate the binding of Oct1 in both untreated and RA-treated P19 cells, but prevents the formation of the ternary complex. The composition of the complexes binding R2-PM-R3 was checked using specific antibodies as indicated on the top. The anti-Oct1 antibody (lanes 6 and 12) strongly inhibits the binding to the oligonucleotide.

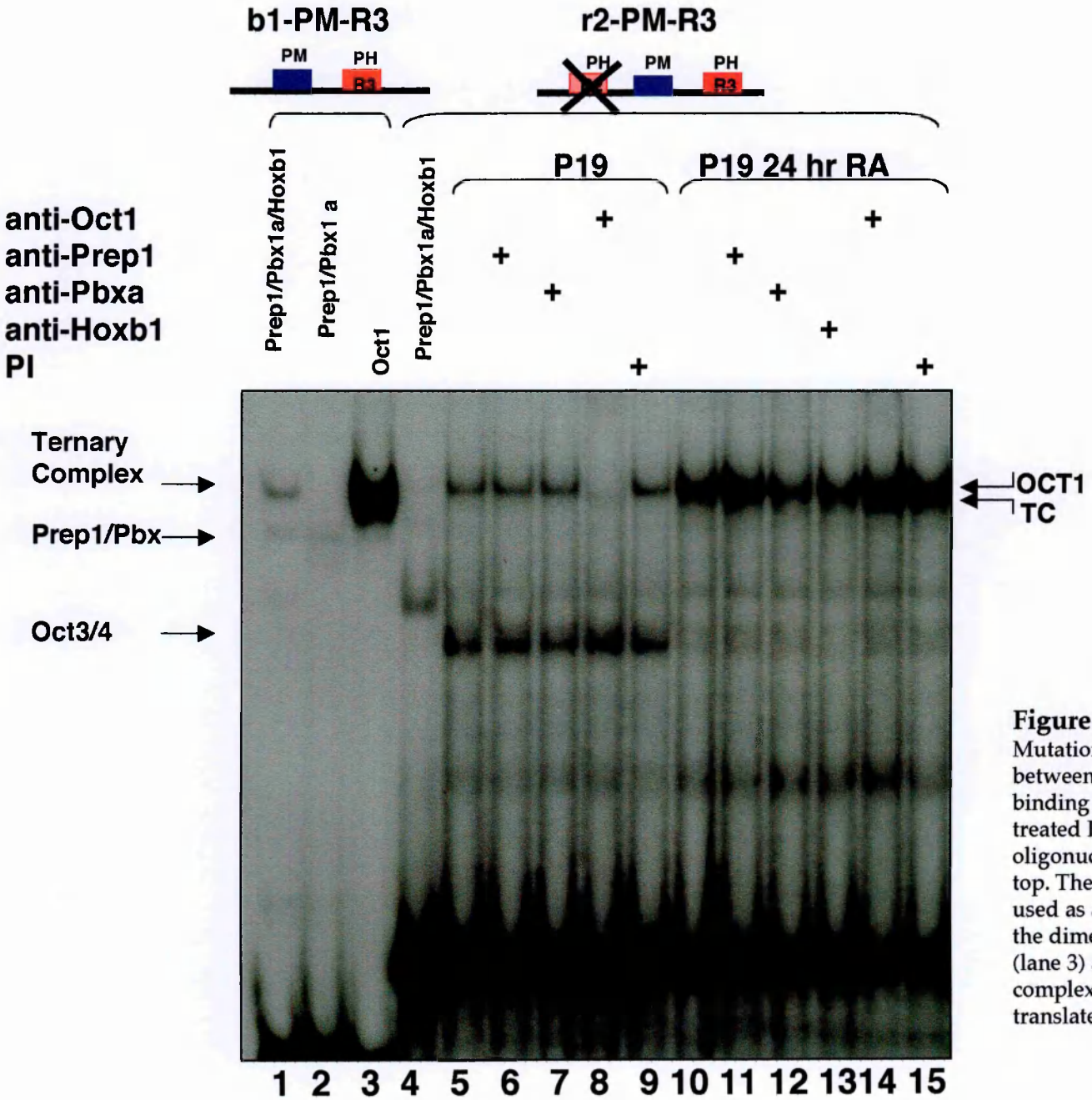


Figure 5.12
Mutations in R2 restore the cooperation between PM and R3 required for the binding of the ternary complex in RA-treated P19 cells extracts. Antibodies and oligonucleotide used are indicated on the top. The b2- PM-R3 oligonucleotide was used as a positive control since it binds the dimer Prep1-Pbx1a (lane2), Oct1 (lane 3) and Prep1-Pbx1a-Hoxb1 ternary complex (TC) formed by *in vitro* translated proteins (lane1) .

As shown in figure 5.10, the complexes formed by b1-PM-R3 oligonucleotide in the presence of untreated cell extracts was completely inhibited by the addition of anti-Oct1 antibodies (lane 4) but not by anti-Prep1, anti-Pbxa or anti-Hoxb1 antibodies (lanes 2, 3 and 5). Moreover, the bands corresponding to the Prep1-Pbx and Pbx/Hoxb1 heterodimers are not present. The absence of Pbx1-Prep1 binding is surprising since their heterodimers are present in the extracts (Ferretti *et al.*, 2000). The absence of Hoxb1-Pbx binding is expected, as Hoxb1 is absent in untreated cells. After RA treatment, which induces *Hoxb1*, the nuclear extract also contained a slower-migrating band, which was only partially inhibited by the addition of either anti-Oct1 (lane 10), anti-Pbx (lane 8), anti-Prep1 (lane 7), or anti-Hoxb1 (lane 9) antibodies. These results suggest that the induction of *Hox genes* results in the formation of a ternary complex. They also show that *Hoxb1* enhancer has a constitutive Oct1 binding activity co-migrating with the ternary complex. In addition a faster band, corresponding to Oct3 binding factor, appears in untreated cells extracts and decreases after RA induction (Di Rocco *et al.*, 2001).

Next, I tested the role of R2 with nuclear extracts from untreated and RA-treated P19 cells in EMSA. The addition of R2 to PM-R3 oligonucleotide revealed the binding of a Prep1-Pbx heterodimer present in untreated and RA treated P19 cells (Fig 5.11, lanes 3, 8). The slower migrating band is again inhibited by the addition of anti-Oct1 antibody both in untreated and treated nuclear extracts (lanes 6 and 12). On the contrary, anti-Prep1 and anti-Pbx antibodies did not inhibit the formation of the slow migrating band in the RA treated extracts respectively (lanes 9 and 10). These results suggest that, as *in vitro*, the R2 site has an inhibitory effect on Prep1-Pbx-Hoxb1 ternary complex formation. Moreover, the re-insertion of R2 restored the ability of Prep1-Pbx heterodimer to bind the oligonucleotide.

Finally, I tested the binding activity of the r2-PM-R3 oligonucleotide. EMSA revealed that the r2-PM-R3 oligonucleotide was no longer able to associate with the Prep1-Pbx heterodimers, suggesting that the Prep1-Pbx heterodimer constitutively binds the R2 site (Fig 5.12). Once again in untreated P19 cell extracts the slow migrating complex was completely inhibited by the addition of anti-Oct1 antibody. On the contrary, anti-Oct1, anti -Prep1, anti-Pbx or anti-Hoxb1 antibodies each partially inhibited the slower migrating bands formed after RA treatment (lanes 11, 12, 13 and 14). Therefore, using nuclear extracts I also observed a cross talk between R2 and the PM-R3 binding motif in the *Hoxb1* enhancer. The presence of R2 inhibited the cooperation between PM and R3 necessary for Prep1-Pbx-Hoxb1 ternary complex formation.

5.6 The *Hoxb1* and *Hoxb2* genes show the same pattern of expression in rhombomere 4 using a different regulatory motif.

To study the *in vivo* behavior of the regulatory sequences analyzed above, I fused the minimal *Hoxb1* and *Hoxb2* sequences (b2-PM-PH and b1-PM-R3) required to bind the ternary complex and the β-actin promoter to the LacZ reporter gene and performed electroporation experiments in the neural tube of live chicken embryos, staining for beta-galactosidase activity.

Figure 5.13 shows that the b2-PM-PH motif is able to direct the restricted expression of the reporter gene in r4 and in migrating neural crest cells (panels

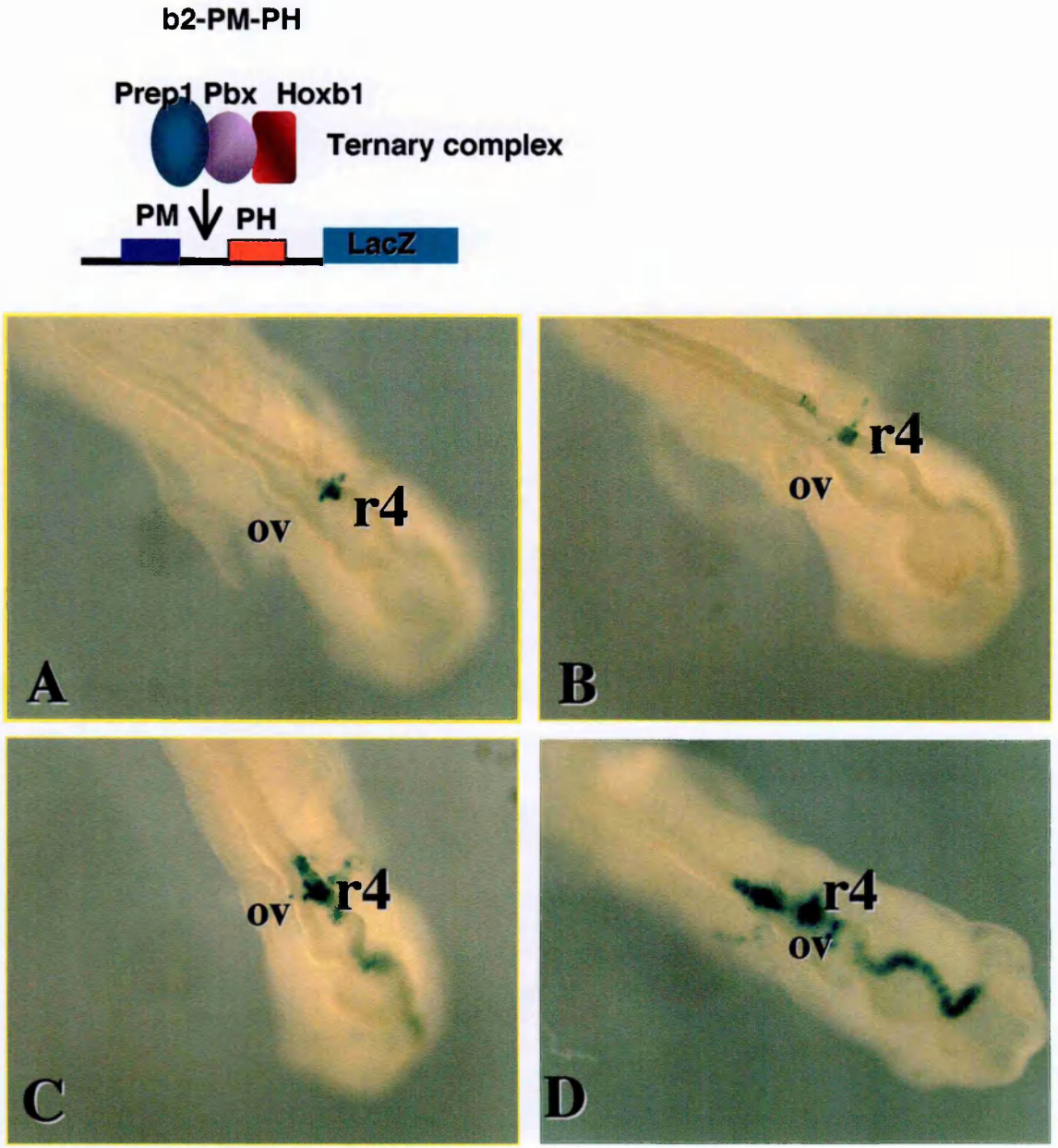


Figure 5.13
The PM-PH motif from the Hoxb2 enhancer drives the expression of the LacZ reporter gene in rhombomere 4 and migrating neural crest cells.
The left side of chick neural tubes were electroporated at 6-7 HH (Hamburger and Hamilton stages) stages with vectors containing the b2-PM-PH (A, B, C and D) fused to the LacZ reporter gene. Beta-galactosidase activity was analyzed at HH16.

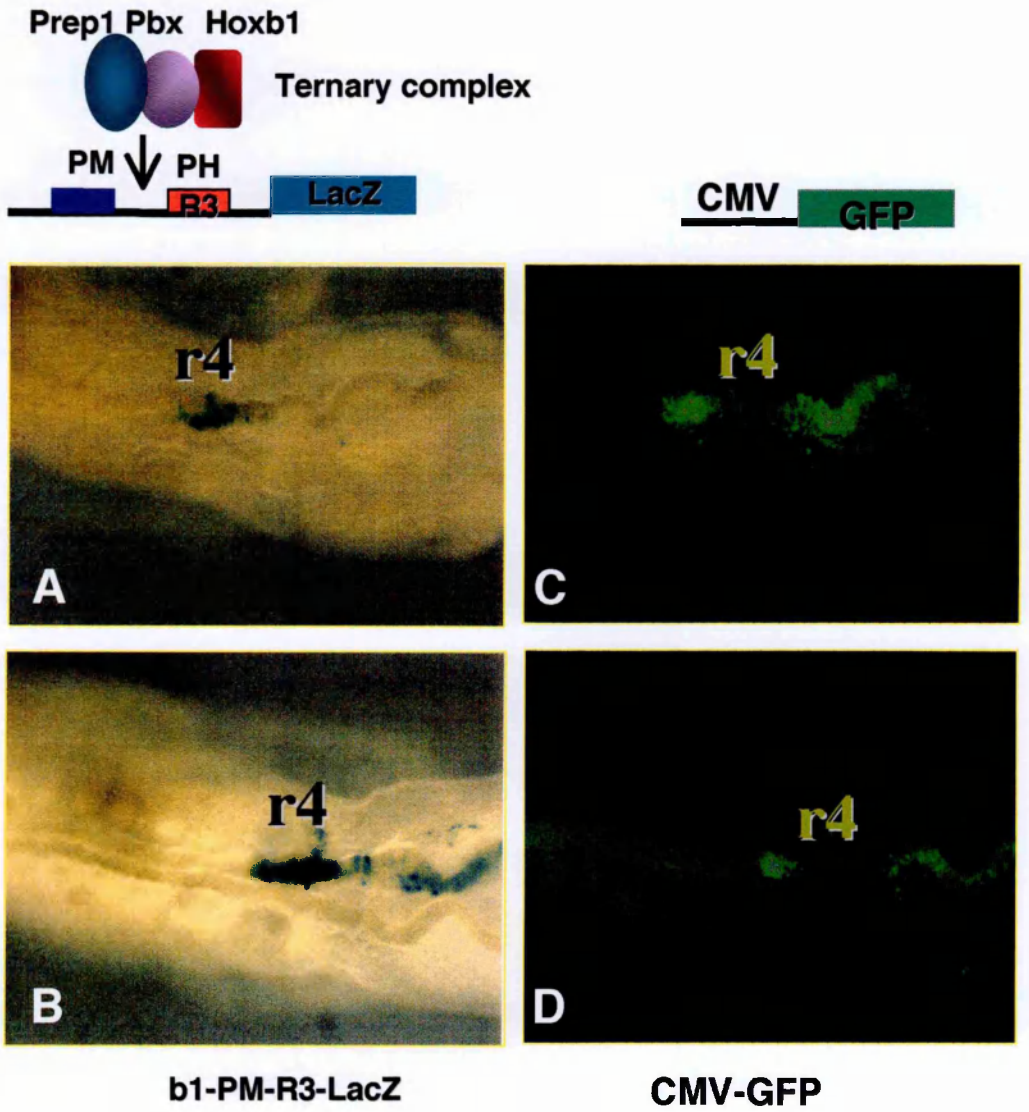


Figure 5.14
The PM - R3 combined element from Hoxb1 enhancer permits lacZ reporter gene expression in rhombomere 4. The PM - R3 oligonucleotide (A and B) drives LacZ expression in r4, while the GFP control vector, driven by the CMV promoter drives the expression of GFP along the entire hindbrain (C and D).

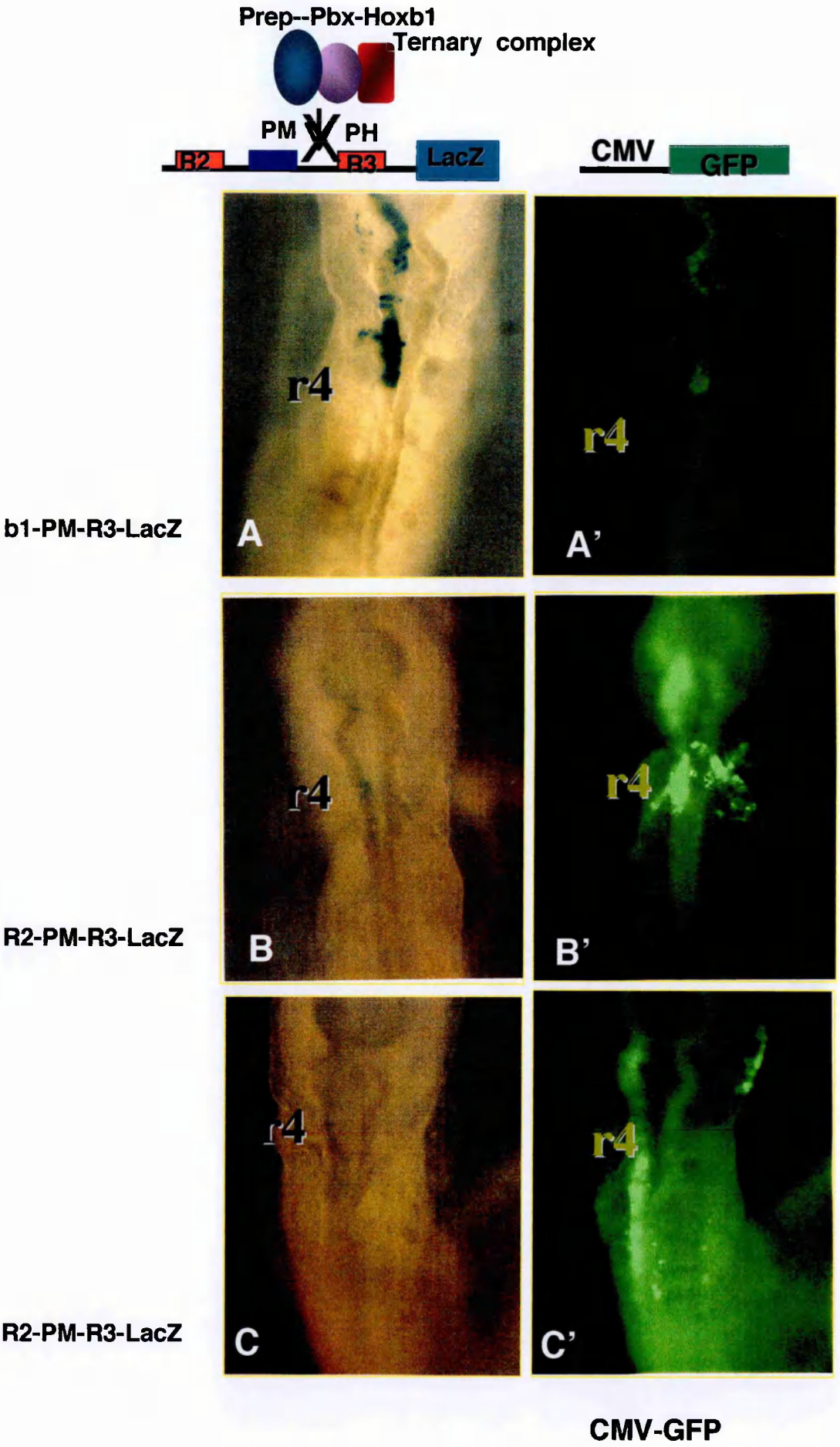


Figure 5.15
Addition of R2 site to PM-R3 combined elements from *Hoxb1* enhancer inhibits promoter activity. The three sets of figurers show three different electroporated embryos. Left side (B and C) R2-PM-R3-LacZ; right side (A′, B′ and C)′ GFP control vector. Embryo in panel A represent the control PM-R3 -LacZ oligonucleotide .

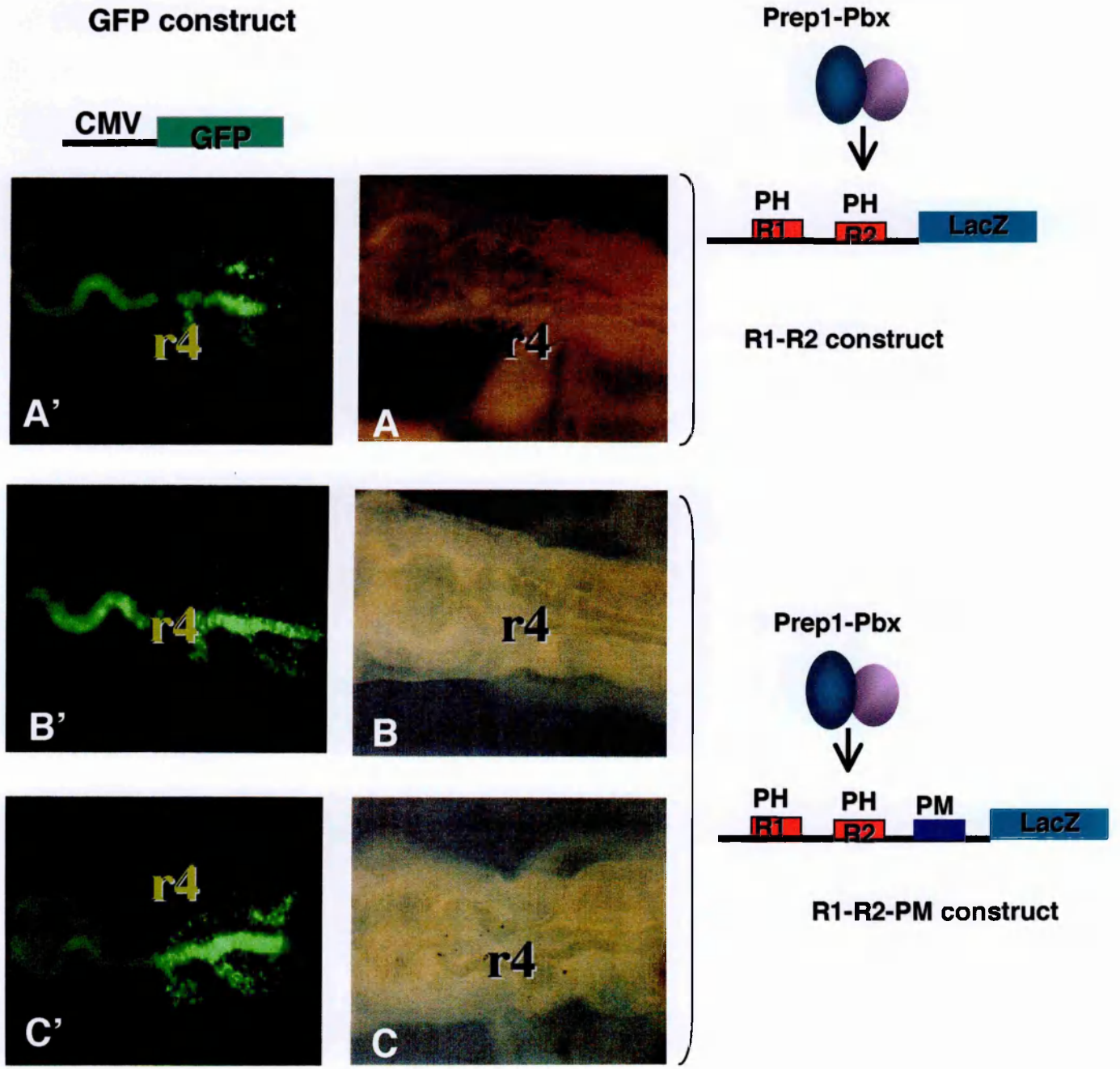
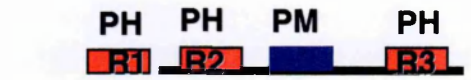


Figure 5.16
R1-R2 and R1-R2-PM motifs from the *Hoxb1* enhancer are not sufficient to permit the expression of the reporter gene in chick r4. The co-electroporation of control GFP vector (panel A′, B′ and C′) demonstrates the efficiency of electroporation. Panel A: R1-R2-LacZ. Panel B and C: R1-R2-PM-LacZ.



Oligo	r4-expression/ tot embryos	%	r4-expression levels	*In vitro* ternary Complex formation
b1-ARE	108/136	80%	+++	N.D.
R1-R2	0/31	0	-	-
R1-R2-PM	0/34	0	-	-
PM-R3	29/52	56%	++	+
R2-PM-R3	20/47	42%	+/-	-

Figure 5.17
Summary of chicken electroporation experiments indicating the structure of the oligonucleotides electroporated, the reporter gene activity and the ternary complex formation *in vitro*.

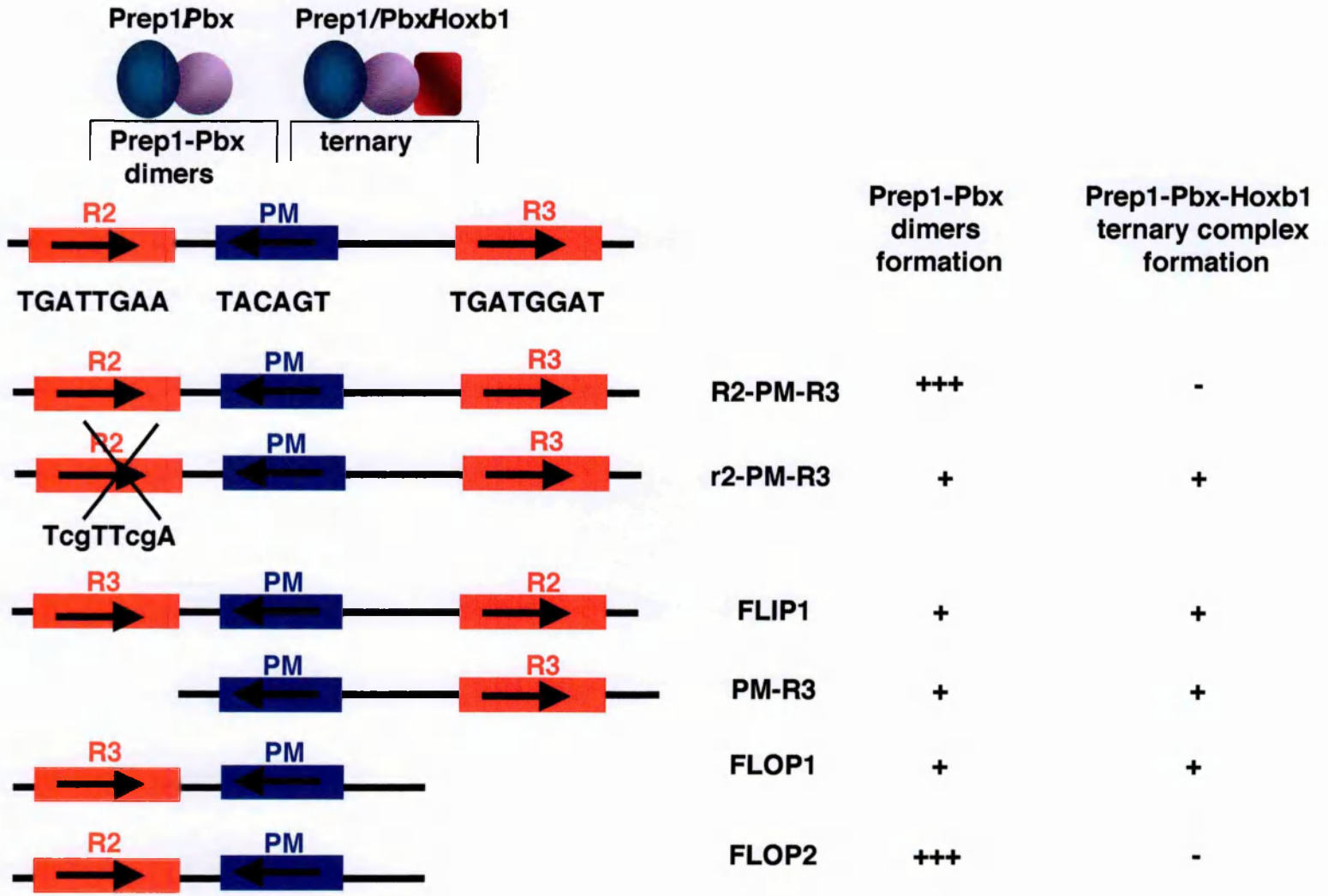


Figure 5.18

Schematic representation of the results obtained in EMSA: these data show that high affinity R2 binding sitr negatively interferes with ternary complex formation.

The addition of R2 toe PM-R3 shifts the affinity of the oligonucleotide from the ternary to the dimeric complex (compare PM-R3 with R2-PM-R3).

The presence of high affinity g of R2 (high affinity biding site for Prep1-Pbx: high affinity PM) probably creates a steric hindrance that prevents the PM-R3 cooperation and thus ternary complex formation (compare R2-PM-R3 with r2-PM-R3). The low affinity Prep/Pbx present in the Flip1 oligonucletide restorse the ability to assemblethe ternary complex.

A-D). The role of b1-PM-R3 motif in expression was tested in the same way using a vector containing GFP as an internal control (Fig 5.14). After co-electroporation of b1-PM-R3-LacZ and CMV-GFP, I found the control GFP expressed along the whole hindbrain length (panels C, D) while the LacZ was only expressed in rhombomere 4 (Fig 5.19, panel A and B).

These results suggest that the combined PM-PH motif, binding the ternary complex, is sufficient to drive *Hox genes* expression in rhombomere 4.

I next tested the *in vivo* role of R2 by adding it to the b1-PM-R3 oligonucleotide. Figure 5.15 shows that lacZ expression driven by R2-PM-R3 is significantly reduced in rhombomere 4 in comparison to the PM-R3 reporter activity (panels A, B, and C). Therefore, inhibition of ternary complex formation results in lower transgene expression in rhombomere 4.

I also investigated the role of R1-R2 tandem PH binding sites by fusing the R1-R2 and R1-R2-PM oligonucleotides to the LacZ reporter gene and using them in chick hindbrain electroporation experiments. Figure 5.16 shows that R1-R2 (panel A) and R1-R2-PM (panel B and C) are not able to mediate the expression of the transgene in rhombomere 4. In fact, no LacZ signal was detected. These results confirm the central role of the PM-R3 site in driving the expression of Hoxb1 in r4 as well as the inhibitory role of the R2 site when combined with PM-R3.

Figure 5.17 summarizes the results of the chicken embryo electroporation experiments together with their statistical significance.

In conclusion, the structure of the *Hoxb1* enhancer appears to be complicated by the presence of multiple PH sites, whose presence modulates the formation of the ternary complex. The inhibitory role of R2 is statistically significant in view of its very low level of expression in r4 in 42% of the embryos which positively stained for LacZ.

Discussion Chapter 5

Mammalian Hox proteins belong to a large family of transcription factors that control cell identity, differentiation and patterning in embryonic development. Members of the same family have the ability to regulate their spatial and temporal expression in the embryos in an autoregulatory manner. Their expression is further regulated by other factors such as morphogene molecules (RA, FGFs) or other transcription factors (for example TALE proteins, Kreisler, and Krox20). The protein players and molecular mechanisms involved in *Hox* regulation are highly conserved from flies to mammals. One of the most studied regulatory mechanisms is the auto-cross-regulatory activity of Hox/PBC complexes. These complexes are able to direct Hox expression in the murine hindbrain in a specific manner.

The Hox/PBC consensus sequence "TGATNNAT", identified as a PH binding site, is found in several *Hox* enhancers (*Hoxb1, Hoxb2, Hoxa3* in mammals and *labial* in Drosophila). It has been demonstrated that the *Exd/labial* binding sites (PH) can have a strong influence on transcription in the Drosophila embryo. Indeed, multimeric (20 bp), PH sites activate specific reporter-gene expression both in Drosophila and in the mouse embryo (Chan *et al.*, 96; Popper *et al.*, 1995).

Moreover, the substitution of two base pairs in each PH element of the multimer modifies its *in vitro* binding specificity and the *in vivo* expression pattern of the transgene in the Drosophila embryo (Chan *et al* 1997). According to this data, the Hox/PBC specificity showed a dependence on the PH sites only. However, a new way of interpretation of this Hox/PBC specificity has been proposed. Li and coauthors have suggested that additional cofactors could determine the specificity of *Hox* response elements in a manner that is largely independent of the binding preferences of Hox/PBC dimers (Li *et al.*, 1999). This group demonstrated that the Exd/labial binding site can alter its specificity if moved into a different context. Thus, the enhancer context seems to exert an important role in determining Hox/PBC expression specificity.

Prep1/Pbx dimers bind to the R2 site and inhibit the cooperation between PM and PH sites in the *Hoxb1* enhancer.

I have previously demonstrated that *in vitro* the Prep1-Pbx-Hoxb1 ternary ternary complex forms on PM+PH combined sites in several *Hox* enhancers and that this sequence, *in vivo*, can have different properties when located in different enhancer contexts (see chapter 4 Fig 4.14 and Fig 4.15). For example, the PM+PH combined site is essential for *r-4-Hoxb2* but not for *r4-Hoxb1* expression suggesting that additional components, like the presence of additional PH sites (R1 and R2) could modulate its activity in *Hoxb1*.

In this study I dissected the *Hoxb1* regulatory element (b1-ARE) and found that R2 negatively interferes with the formation of the Prep1-Pbx-Hoxb1 ternary complex on the combined PM-R3 site (Fig 5.2). Indeed, PM-R3, but not R2-PM-R3, permits the assembly of the ternary complex (Fig 5.2, lanes 2 and 5). Interestingly, R2-PM-R3 binds Prep-Pbx dimers with much higher affinity than PM-R3 (Fig 5.3, lane 2).

The inhibitory effect of R2 on ternary complex formation is further demonstrated by the observation that mutations in R2 restore the ability of the PM-R3 motif to assemble the ternary complex (Fig 5.4, lane 11). These mutations also strongly affect the Prep1-Pbx binding demonstrating that R2 represent the high affinity binding site for the Prep-Pbx dimer (Fig 5.4 lane compare lanes 3 and 9). This strong binding depends mainly on R2 as the mutated R2-pm-R3 oligonucleotide, containing mutations in PM, (Fig 5.2 lane 7) and the R2-PM motif, missing R3, continue to strongly interact with Prep-Pbx dimers (Fig 5.3, lane 2). Moreover, R2, not only has an inhibitory effect on PM-R3 binding activity, it also does not substitute for R3 since R2-PM is not able to form a Prep1-Pbx-Hoxb1 ternary complex (Fig 5.3, lane 3).

The comparison of the R2 binding sequence and PM sequence revealed the presence of 3 bases (CTC) flanking the consensus sequence, which I have demonstrated to be essential for the high affinity of the PM site (see Fig 4.11 and Fig 7.1). This suggests that R2 is a high affinity binding site for Prep1-Pbx.

R2-PM and PM-R3 have different binding properties *in vitro*. Differences between the R2-PM and PM-R3 motifs that could explain this data are: the distance between PM and PH sites, their relative position and the high binding affinity of R2 for Prep1-Pbx dimers. The data presented in this chapter show that the distance and the relative position of R3 to PM do not alter the ability of R3 and PM to cooperate in the ternary complex assembly (Fig 5.7, line 12). These

data are in agreement with previous data in which orientation and spacing between PM and PH has shown to be very flexible (Fig 4.15; Chan *et al.*, 1999). When I moved R2 downstream of PM, Flop1, R2 was no longer able to bind Prep-Pbx dimers with high affinity nor substitute for R3 in the formation of the ternary complex (Fig 5.7, compare lanes 3 and 9). Flipping R2 and R3 (R2-PM-R3: Flip1) did not affect the ternary complex (Fig 5.8, lane 6). Overall these data indicate that the relative position of R2 is important for its function and that the high affinity binding of Prep1-Pbx to R2 is responsible for the inhibition of ternary complex formation.

The fact that when R2 is located upstream of PM binds Prep1-Pbx with strong affinity allowed me to hypothesize a model in which the presence of Prep1-Pbx dimers bound to R2 creates a steric hindrance that prevents ternary complex formation. Indeed, ternary complex formation is rescued when the dimer binding is reduced (Fig 5.17). In addition, the high amount of Prep1-Pbx dimer forming on R2 may prevent the interaction of these two factors with Hoxb1 and hence formation of ternary complex. In conclusion, the data demonstrated that R2-PM-R3 preferentially binds the dimer rather than the ternary complex.

Induction of Hoxb1 in RA-treated P19 cells changes the architecture of the *Hoxb1* enhancer.

In order to investigate what happens in a more physiological context, I used murine embryonic carcinoma cells (P19) which retain most of the characteristics of primitive neuroectodermal cells and in which it is possible to induce the Hox expression on retinoic acid (RA) treatment. Nuclear extracts purified from RA-treated P19 cells form the Prep1-Pbx-Hoxb1 ternary complex on PM-R3 but not on R2-PM-R3 (compare Fig 5.10, lane10 and Fig 5.11, lane 12). Once again mutations in R2 restore the formation of the ternary complex (Fig 5.12 lane 14). Surprisingly, both in treated and untreated cells, Prep-Pbx complexes were not able to bind the PM-R3 oligonucleotide, while the addition of R2 conferred this ability in both uninduced and induced P19 cells (Fig 5.11 lanes 3 and 8). This suggests that Prep1-Pbx dimers specifically bind R2, as mutations in R2 inhibit the binding of this complex thus reproducing the pattern seen *in vitro* with translated proteins (Fig 5.8). This data are in agreement with R3 being a much lower affinity binding site for Prep1/Pbx1a.

In addition to the three repeats, an octamer-like sequence ATGCAAAT) located between R2 and R3 has been characterized (Di Rocco *et al.*, 2001). The octamer sequence binds transcription factors belonging to the POU family (Ryan and Rosenfeld, 1997). These factors are characterized by the presence of the POU domain, a bipartite DNA binding structure containing an N-terminal POU specific region (POUs) and a POU homeodomain (POU-HD) connected by a linker region. POU proteins can bind the DNA as homo- and heterodimers. In addition to their DNA binding functions, the POU domain can participate in protein-protein interactions with other transcriptional regulators.

Using P19 nuclear extracts, I showed that the main factor binding to the Hoxb1 oligonucleotides is Oct1, and that while Oct1 binding does not change after RA treatment, Oct3 binding decreases. These data suggest that a modification in either the concentration of Oct3 or the enhancer conformation

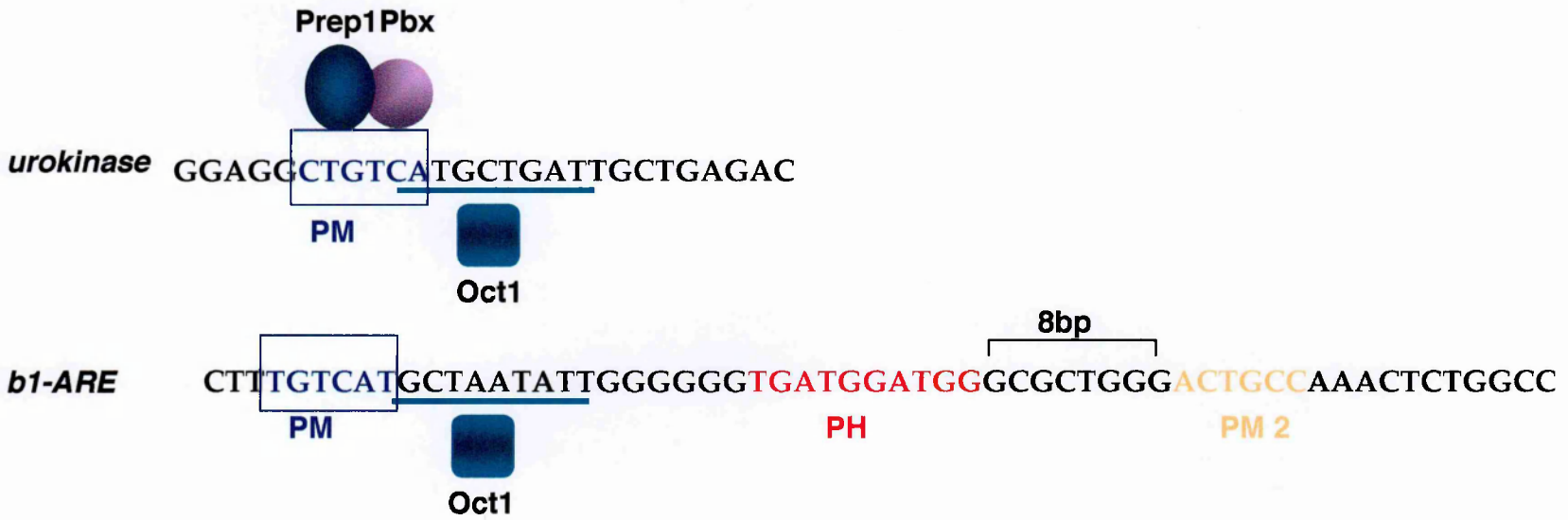


Figure 5.19
Schematic representation of the *urokinase* and *Hoxb*1 enhancers showing the overlapping binding sites for Prep1-Pbx (boxed in blue) and Oct1 (underlined in green).
The orange PM binding site 8 bp downstream to R3 represents a putative Prep/Pbx site (PM2), probably able to cooperate with R3 in ternary complex formation.

occurs after RA treatment and Hox induction (Fig 5.10, Fig 5.11, and Fig 5.12 compare control and treated cells).

Two observations are important in proposing a model to explain *Hoxb1* regulation. Firstly, Oct1 has been shown to act as a repressor in concert with other transcription factors including C/EBP and Pbx proteins (Saleh *et al.*, 200). Secondly, regulators of Pbx/Hox dimer formation are members of the histone acetylase (HAT) and histone deacetylase families (HDAC), which modify chromatin, as well as non histone proteins (Shen *et al.*, 2001). It is well known that the acetylation status of chromatin modulates the accessibility of transcription factors to the regulatory regions.

Thus, a possible model could involve Prep1, Pbx, Oct1 and Oct3 as repressor factors able to recruit HDAC to the regulatory element (*ARE*) of *Hoxb1* in uninduced cells. After RA stimulation, the architecture of *b1-ARE* changes and binding of Oct3 decreases, leaving the region accessible to Hoxb1. Thus, Prep/Pbx dimers can act as repressors in the absence of Hoxb1 (the RA un-induced condition).

All these data suggest that the modulation of protein-protein interactions, rather then the more elusive protein-DNA interactions, provides a strategy to modulate transcriptional processes during embryonic development.
Understanding the exact mechanism by which the b1-ARE element is able to modulate *Hoxb1* expression *in vivo* is a difficult task. However a possible model should involve the modulation of the chromatin status and the PH/PM sites could be required to open and close the chromatin in this region.

An enhancer structure similar to that of *Hoxb1-ARE* has been described for the *urokinase* enhancer. Indeed in both enhancers Prep1-Pbx and Oct1 are present in an overlapping binding site (Fig 5.19). For the *urokinase* enhancer it has been demonstrated that Oct1 is involved in the regulation of basal transcriptional activity (Palazzolo *et al.*, 2000). Furthermore, reporter gene expression of transgenic mouse embryos carrying a LacZ reporter gene fused to the mutated Oct binding site of *Hoxb1* does not impair the ability of the r4 enhancer to direct the expression in r4. However, extending the mutation to the flanking region containing the PM site, influences the response to RA but not r4-restricted expression (Di Rocco *et al.*, 2001).

PM-PH elements are instructive sequences *in vivo*.

I have demonstrated that the PM-PH sequences from *Hoxb1* and *Hoxb2* enhancers are both able to assemble a Prep1-Pbx1a-Hoxb1 ternary complex *in vitro* (Fig 4.3) and they are also sufficient to drive the r4-restricted expression of a reporter gene in the electroporated chicken hindbrain (Fig 5.13 and 5.14). The PM-PH motifs of *Hoxb1* and *Hoxb2* are substantially different in terms of their sequences, the spacing between the two sites and their binding of additional proteins. However, they have the same effect *in vivo*.

Furthermore, the presence of R2, as *in vitro* experiments, exerts a negative effect on PM-R3 cooperation and as a consequence inhibits reporter gene expression *in vivo*. Due to its large size the R1-R2-PM-R3 (80 bp) is difficult be tested by EMSA and therefore I still do not know the role played by R1 in ternary complex formation.

These data prove a direct correlation between ternary complex formation and r4-restricted expression. Moreover they suggest that the *Hoxb1* enhancer (b1-ARE) may contain additional components promoting the expression in r4 and able to overcome the negative effect of R2. Among these components are: the multiple PH sites, the presence of additional transcriptional factor binding sites such as Oct1/Oct3 and the presence of an additional PM site downstream to R3, the role of which is under investigation. The additional PM site is located 8 bp (PM2) downstream of R3 and it could cooperate with R3 to form a ternary complex even in the presence of high levels of Prep-Pbx dimers bound to R2 (Fig 5.19).

The data presented in this chapter includes data obtained in collaboration with Robb Krumlauf's group and Paco Cambronero. The work with the chicken transgenic embyos was done at the Stowers Institute for Medical Research, Kansas City, USA.

CHAPTER 6
RESULTS AND DISCUSSION IV

Inactivation of *Prep1* gene

6.1 Inactivation of Prep1 in murine ES

In order to study the *in vivo* function of Prep-1 during mouse embryonic development, I decided to generate mice null for the *Prep1* murine gene. Random integration of retroviral enhancer-less vectors in ES cells can be used to interrupt and inactivate genes. This strategy is termed *gene trapping* and has been used to inactivate a large number of genes in ES cells. All the sequences of trapped genes are collected in the public data banks (Fig 6.1; Brian *et al.*, 1998). I searched the Lexicon genetics data bank of integration sites using the mouse Prep1 cDNA sequence and found several clones containing the insertion of a retroviral targeting vector in the *Prep1* gene. I chose clone OST7871 in which the integration of the targeting vector occurred 2800 bp before Prep1 ATG. Figure 6.1 shows the targeting strategy and the integration site in the OST7871 clone that interrupts the murine *Prep1* gene (Fig 6.1). The insertion of the Lexicon genetics retroviral vector occurs into exone one (Fig 6.1). After the integration, two fusion transcripts should be obtained. One contains the puromycin-resistance sequence (pur) and the 5'UTR and first exon of *Prep1* and its driven by the ubiquitous *Pgk* promoter (Fig 6.1). The second contains the 5' UTR sequences of *Prep1* and the neomycin. The presence of multiple stop codons between puromycine and Prep1 cDNA abolishes Prep1 translation.

The expected wild type and mutated *Prep1* alleles were tested by Southern blot analysis of DNA extracted from targeted ES cells (Fig 6.2). Because both the targeting vector sequence and the precise sequence of genomic *Prep1* DNA were not available I first set up screening strategy by Southern blot analysis to discriminate between wild type and mutated alleles. The comparison between genomic DNA obtained from wild type and heterozygous cells digested with different restriction enzymes showed EcoRI as the only enzyme able to give an informative restriction pattern. The fragment encompassing the critical EcoRI site, was used as a probe in the Southern blot analysis. Because of the presence of an EcoRI site in the third intron, probe 2-3 hybridizes to two bands from the wild type allele (12 kb and 8 kb). The introduction of an additional EcoRI site by the targeting vector results in the shortening of the 8 kb band down to a 6 kb band, thus allowing the discrimination between wild type and null allele (Fig 6.2 compare lane 1 with 2 and 3). The heterozygous DNA shows all three 12 kb, 8 kb and 6 kb bands.

After receiving heterozygous mice from Lexicon Genetics I crossed them and generated viable homozygous knockout mice. I verified the absence of functional Prep1 protein by immunoblotting analysis on nuclear extracts comparing wild-type and knockout mice tissues.

Figure 6.3 shows the Western blot analysis of thymus, brain, spleen and liver nuclear extracts. Homozygous mice express no full-length Prep1 protein.

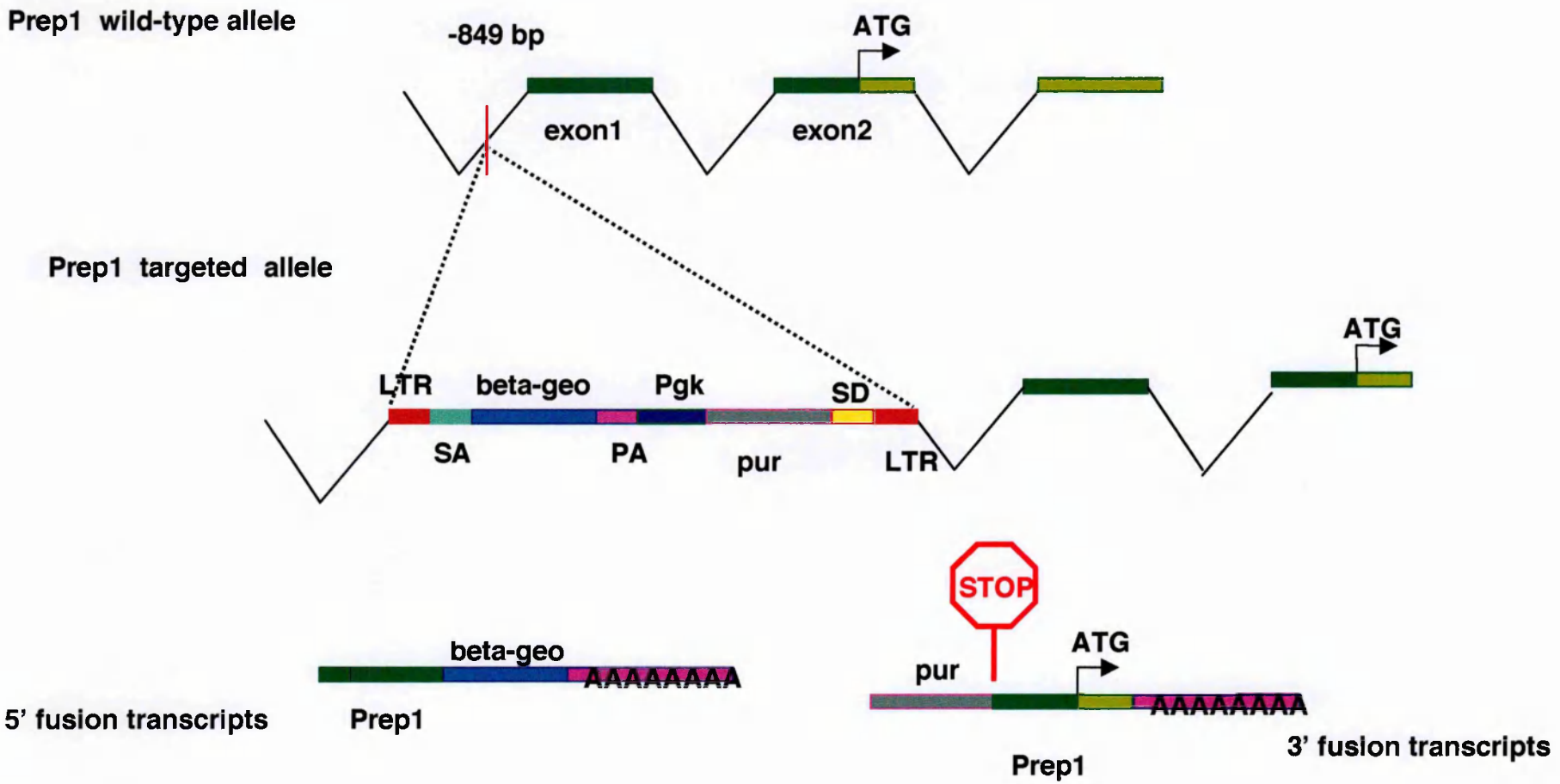


Figure 6.1

Schematic representation of the gene trapping approach used to inactivate the Prep1 gene. The targeting vector contains a promoter (*Pgk*) that works in ES cells and directs the expression of selective marker gene puromycone (pur). The selective marker gene lacks a polyadenylation sequence (PA) and instead has a splice donor sequence (SD). The resulting fusion transcript allows for the direct sequencing of 3′ RACE products after PCR amplification (Brian et al., 1998).

The targeting vector contains long terminal repeats (LTR) the splice acceptor sequence (SA), the fusion cDNA LacZ/ neomycine (beta-neo). The insertion of the targeting vector maps to the introne (at -849 bp before the first exon) at -2800 bp from the ATG. Integration of the two transcripts results in a 5′ fusion transcript and a 3′ fusion transcript. The presence of multiple stop codons between puromycine and Prep1 cDNA aboloshes Prep1 translation.

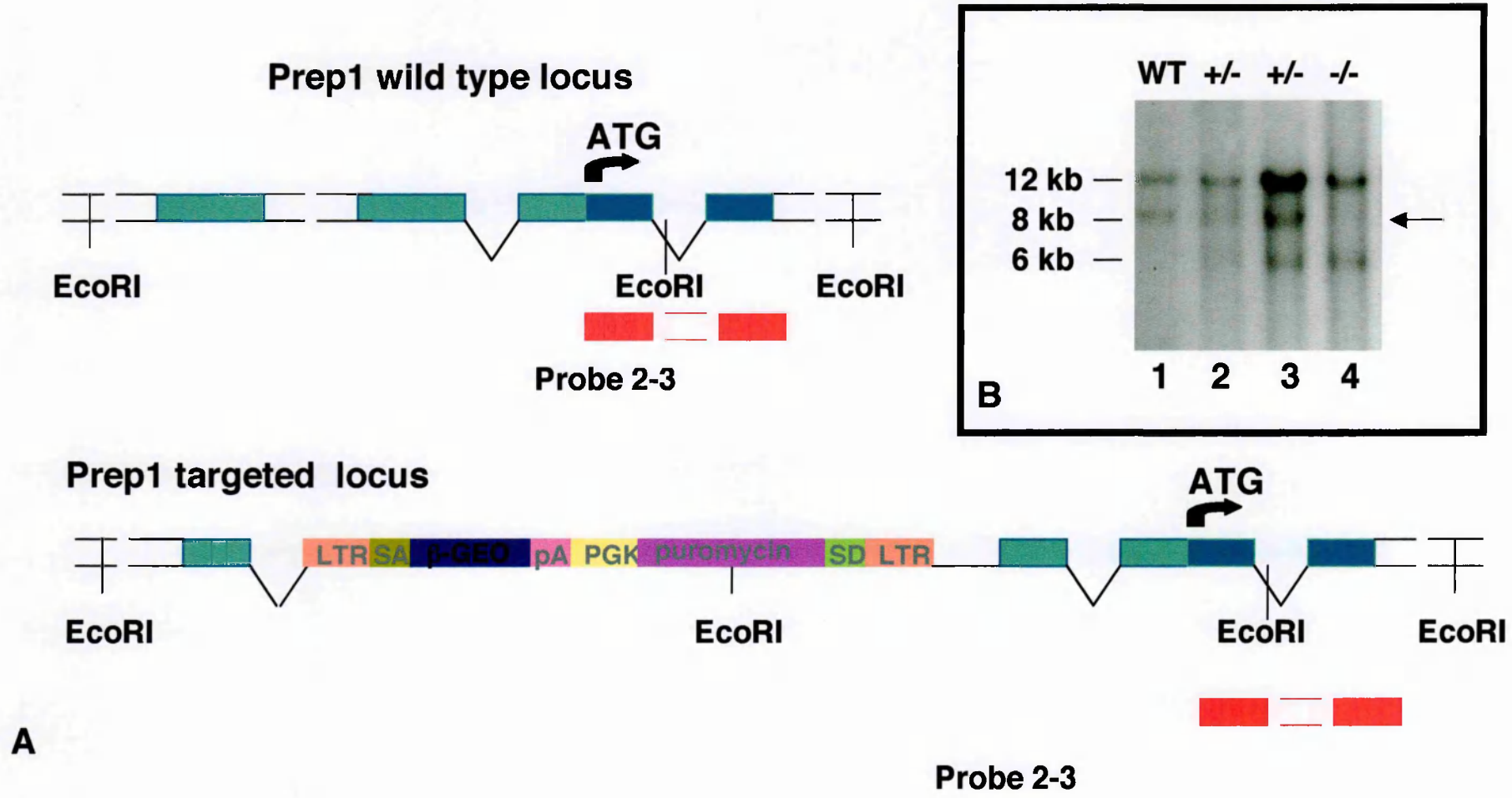


Figure 6.2
Targeted inactivation of *Prep1* gene.
A) Restriction map of genomic DNA containing the 5'region of the *Prep1* gene with the first three exons (green boxes). The red bars illustrate the probe used for Southern blot analysis. Probe 2-3 contains 132 bp of Prep1 cDNA after the ATG site and detected 12 and 8 kb bands in WT cells and 12 , 8 and 6 kd bands in mutated cells.
B) Digestion with EcoRI gives two bands in the WT (12 and 8 kb; lane 1), three bands in heterozygous (12, 8, and 6 kb; lane 2 and 3) and two bands in the homozygous (8 and 6 kb: lane 4).

Moreover the absence of atypical bands in nuclear extracts from knockout tissues suggests that no truncated proteins are produced. Finally, the presence of a slower migrating band in brain tissues may represent either a truncated product of Prep1 or a cross-reaction of anti-Prep1 antibody with other members of the Prep family, such as Prep2 (Figure 6.3, lanes 6 and 8). Prep1+/- mice express half of the normal levels of Prep1 (Fig 6.2, lanes 3, 8, 10 and 13).

6.2 Prep1 null mice phenotype depends on the genetic background.

Homozygous Prep1 null mice are viable and fertile and displayed no apparent abnormalities. Inactivation of the *Prep1* gene was performed in 129/SvEvBrd embryonic stem cells (ES). Targeted ES clones were injected into C57BL/6 albino host blastocysts. Lexicon Genetics provided heterozygous mice having a 50% 129/SvEvBrd and 50% C57BL/6 albino mixed genetic background. The F1xF1 cross gave a percentage of homozygous knockout mice consistent with the expected Mendelian frequencies. The resulting knockout mice did not show any evident phenotype.

However, by inter-crossing the next generations (F2XF2 and F3XF3) of heterozygous mice we observed a percentage of Prep1-/- pups much lower than expected. These data suggest that knocking-out Prep1 may give an embryonic lethal phenotype and that the Prep1-/- phenotype is probably dependent on the genetic background. For this reason I decided to clean the genetic background of our mice by backcrossing them with C57Bl/6J wild type mice. In order to obtain an isogenic C57Bl/6J line is necessary to backcross mice for at least 9 generation. I am currently obtaining and analyzing the 6 th generation of these mice.

6.3 Prep1 is essential for fetal development

The potential embryonic lethal phenotype in Prep1-/- mice is also strongly supported by the observations that Prep proteins are able to control the sub-cellular localization and stabilization of Pbx proteins and that both Pbx1 and Pbx3 knockout mice show an embryonic lethal phenotype (Selleri personal communication and Selleri *et al.*, 2000). In order to determine the specific time of embryonic lethality I started to analyze embryos at different stages of development obtained by intercrossing heterozygous mice from the 5 th generation.

I observed that prior to E12.5 d.p.c. Prep1-/- embryos could not be morphologically distinguished from the wild type or heterozygous embryos. However, at 14.5/15.5 d.p.c. Prep1-/- embryos were consistently smaller, showing retardation in the embryonic development (Fig 6.4). Moreover they exhibit massive subcutaneous edema, generalized pallor, diminished vascularization, smaller livers and abnormal orientation of forelimbs (Fig 6.4). Although more severe, these abnormalities have been observed in Pbx1-/- embryos. Moreover approximately 25 % of knockout Prep1-/- embryos show abnormal eyes or the complete absence of eyes (data not shown).

Since it is known that Pbx2/Prep1 dimer is able to bind the Pax6 regulatory region (Mikkola *et al.*, 2002) we could hypotesize a role for Prep1 in eye development.

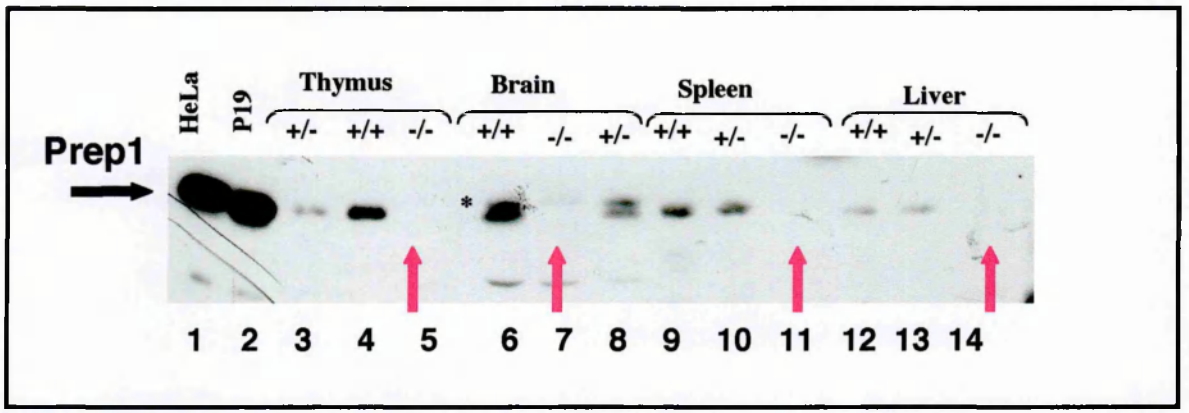
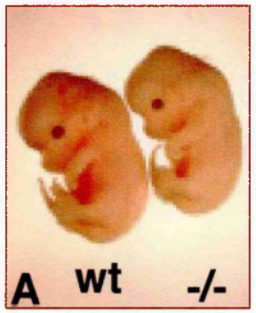


Figure 6.3

Prep1 protein is absent in Prep1-/- mice tissues. Nuclear protein extracts obtained from wild-type (+/+), heterozygous (+/-) and homozygous (-/-) murine tissues (as indicated on the top) were analyzed by Western blot using the anti-Prep1 antibody. Genotypes determined by Southern blot are listed at the top. Red arrows indicate the absence of Prep1 protein in the thymus, brain, spleen and liver nuclear extract purified from Prep1-/- mice. The control lane 1 and 2, contain nuclear extract purified from HeLa and P19 cell lines, respectively. The brain nuclear extracts contain double band indicates with * may correspont to Prep2 .



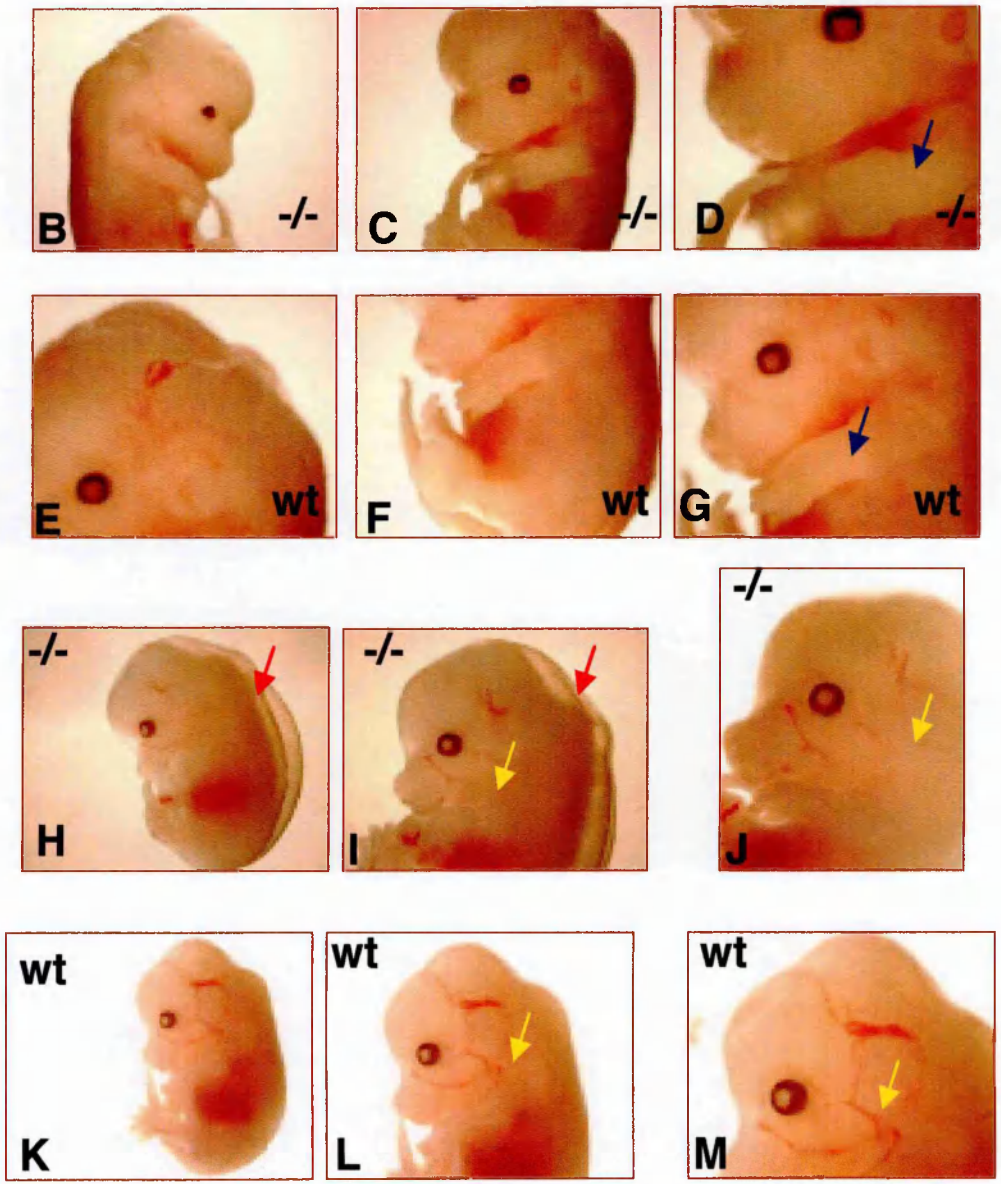


Figure 6.4
Gross morphology analysis of wild type and Prep1-/- embryos at 15.5 d.pc. (A, B, C, D, E, F and G) and 14.5 d.p.c. (H, I, L, J, K, L and M). Mutant mice show smaller dimensions (A), massive subcutaneous edema (H, L red arrow), pallor, reduced vascolarization (compare I and J with L and M; yellow arrow), abnormalities in limb bud development (blue arrow).

I observed a variable penetrance of this phenotype probably due to the mixed genetic background. Furthermore I need to consider an intrinsic incomplete penetrance of the Prep1-/- phenotype.

These results are very preliminary and need further investigation to better understand the Prep1-/- phenotype, preferentially using mice from a congenic genetic backgrounds.

6.4 Inactivation of the *Prep1* gene affects the PBC protein level both in mouse and in zebrafish.

Several data have shown that the interaction of Meinox proteins with PBC has several molecular consequences. Firstly, it allows the nuclear localization of these proteins as they have no nuclear localization signal. Secondly, the interaction with Meinox proteins prevents the nuclear export of PBC, thereby allowing interaction with the Hox proteins (Berthelsen *et al.*, 1999; Abu-Shaar *et al.*, 1999). In addition, Meinox proteins have an important regulatory function in stabilizing PBC proteins. For example, in *D. Melanogaster*, Hth is required to maintain the level of Exd (Kurant *et al.*, 2001). In chicken and zebrafish embryos dominant negative Meis1 and Meis3 overexpression reduces the level of Pbx4/*lazarus* (Mercader *et al.*, 1999; Capdevila and Belmonte, 1999; Waskiewicz *et al.*, 2001). In order to further evaluate the stabilization properties of Prep-1 *in vivo* we decided to analyze the levels of the Pbx proteins, in two organisms in which Prep1 had been inactivated, mouse and zebrafish.

As shown in figure 6.5 inactivation of Prep1 in mice results in the selective reduction of Pbx protein forms in a tissue-dependent manner. The level of Pbx1a decreases in thymus, lung, and cerebellum nuclear extracts (Fig 6.5A). Pbx2 levels strongly decrease in thymus, spleen and lung (Fig 6.5B). The expression of Pbx3 slightly decreases in cerebellum (Fig 6.5C). Furthermore, Pbx4 that is normally expressed in testis, is totally absent in Prep1-/- (Fig 6.5D).
All these observations demonstrate that Prep1 the absence of Prep1 results in decreasing of specific Pbx members from the nuclei of different tissues.

In zebrafish the homolog of mammalian Prep1 is termed *prep1.1*. As with mammalian Prep1, *prep1.1* is also maternally expressed in zebrafish (Waskiewiscz *et al.*, 2001; Choe *et al.*, 2002; De Florian *et al.*, 2003, submitted). Moreover, at 24 hour post fertilization (hpf) *prep1.1* is ubiquitously expressed, but at 72 hpf it becomes restricted to the head, brain and otic vesicle. The *prep1.1* gene was inactivated using *prep1.1* antisense morpholinos injected into zebrafish embryos (De Florian *et al.*, 2003, submitted). The phenotype of morphants was characterized by abnormal head development, thinner yolk extension, anomalous tail curvature, atrophic pectoral fins, and abnormal melanocyte distribution and embryonic lethality.

A similar phenotype is observed in a mutant zebrafish called *lazarus*, and results from mutation of the zebrafish Pbx4 gene (notice that the number of Pbx genes in mouse and in zebrafish is different). These results suggest a concerted action of Pbx-Prep proteins during zebrafish development. In order to test if the inactivation of zebrafish *prep1.1* affected the levels and localization of Pbx proteins I performed immunoblotting experiments using the pan-Pbx antiserum to compare the levels of Pbx proteins in nuclear and cytoplasmic extracts

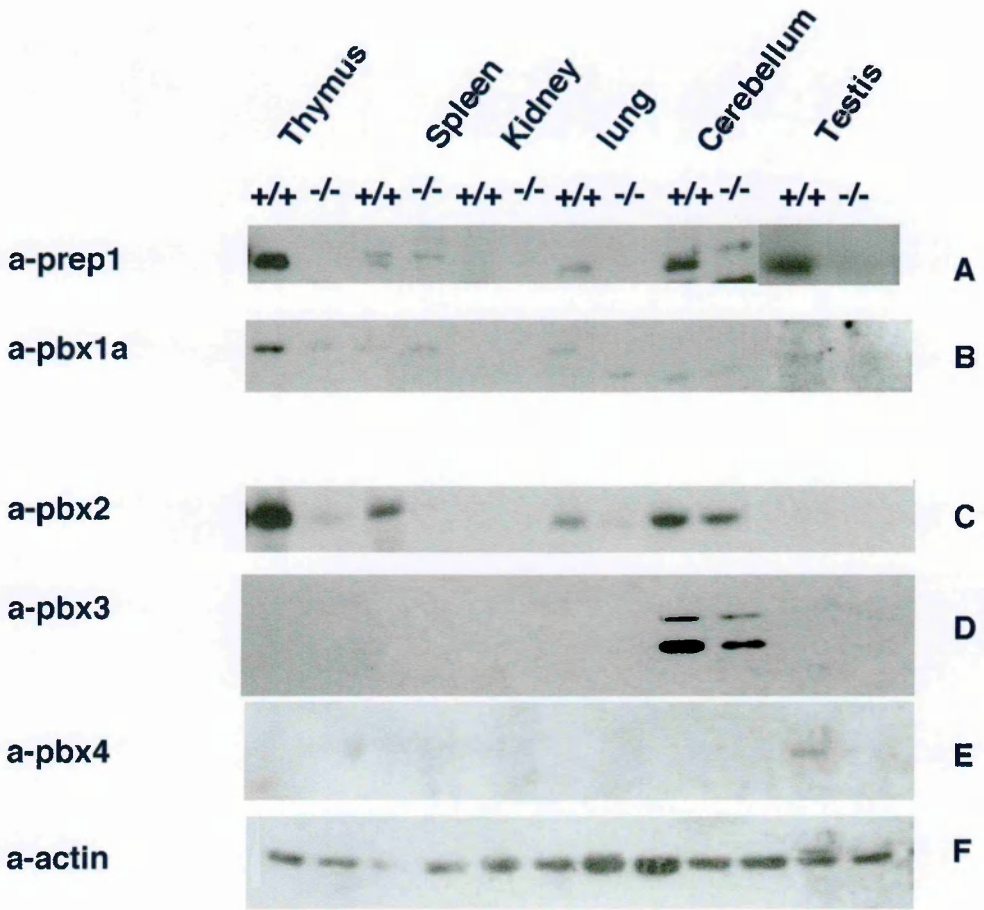


Figure 6.5

Analysis of Pbx protein levels in Prep1-/- mice tissues. Inactivation of prep1 results in reduction of pbx proteins in a tissue dependent manner.

Nuclear protein extracts obtained from wild-type (+/+), and homozygous (-/-) murine tissues (as indicated on the top) were analyzed by Western blot analyzing the anti-Prep1 (A), anti-Pbx1 (B), antiPbx2 C) , anti -Pbx3 (D), anti-Pbx4 (E) and as an internal control, anti-actin antibodies (F). Genotypes determined by Southern blot are listed at the top (Western blot analysis was performed by Longobardi E.).

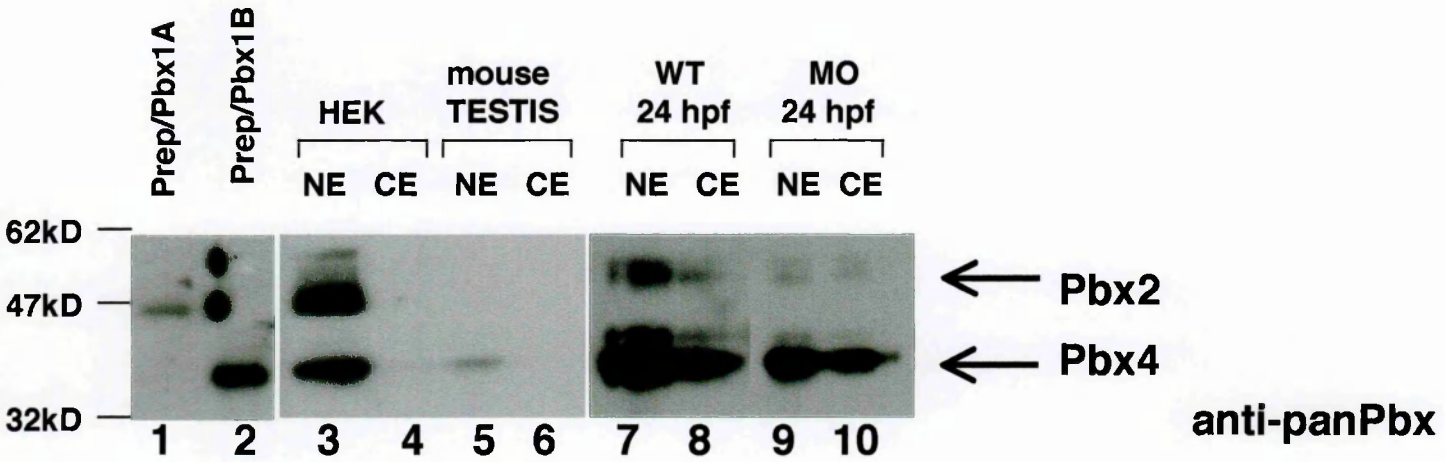


Figure 6.6
Immunoblotting analysis of the levels of Pbx proteins in wild type and *prep1.1* morphants embryos. Pan-Pbx antibody was used to detected Pbx protein levels in nuclear (NE) and cytoplasmic (CE) extracts obtained from wild-type and *prep1.1* morphants 24 hpf zebrafish embryos.
The controls nuclear extracts lanes 3 and 6 contains proteins purified from human HEK293 cells (and from mice testis respectively.
In vitro translated Pbx1a and Pbx1 b proteins were used as control of the molecular weight (lanes 1 and 2).

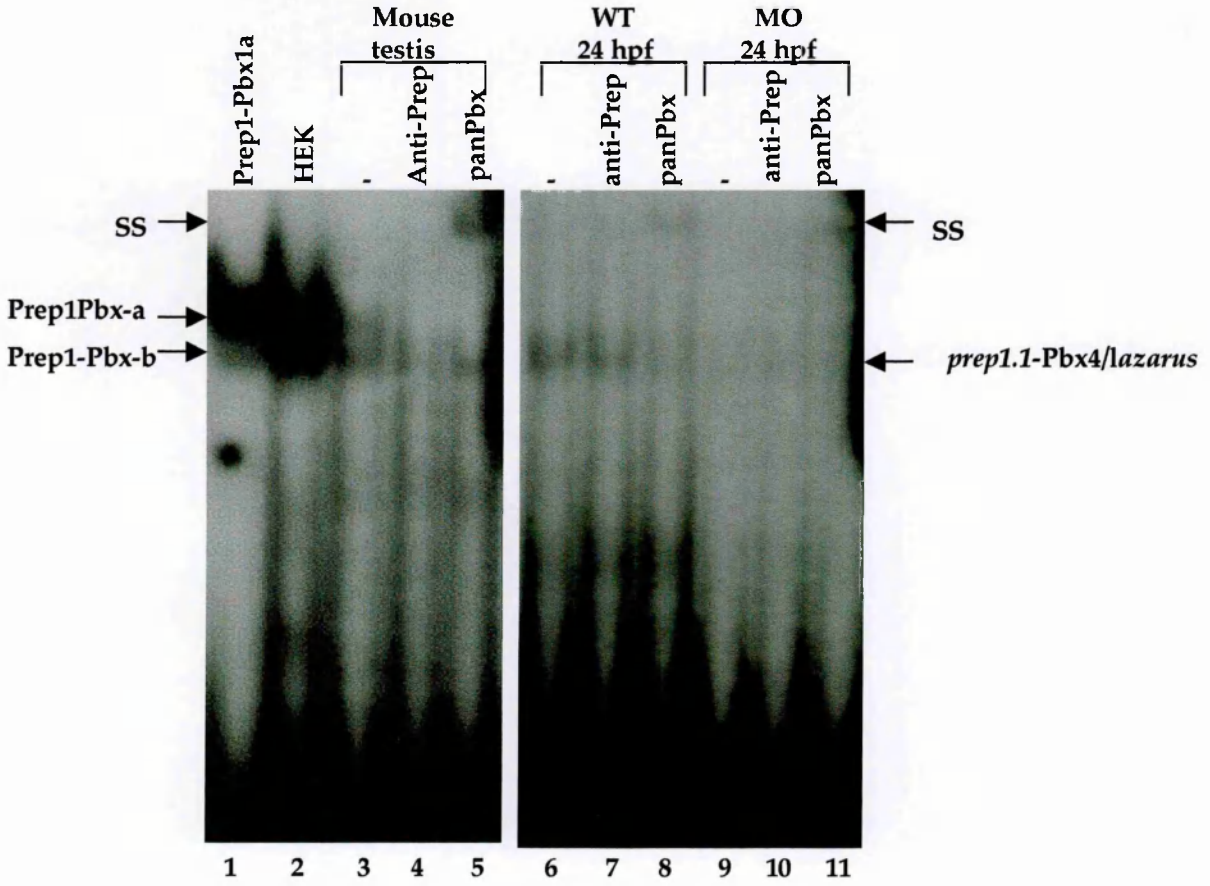


Figure 6.7

EMSA using the O-1 oligonucleotide with nuclear extracts purified from wild type (WT) and *prep1.1* morphant 24 h.p.f. embryos (MO). Prep1-Pbx complexes present in nuclear extracts purified from murine testis bind the O1 oligonucleotide(lane 3), the addition of pan-Pbx antibody produced a supershifted (SS).

A band binding the O-1 oligonucleotide corresponding to *prep1.1* -Pbx4/*lazarus* was detected in WT(lane6) but not in *prep1.1* morphants 24 h.p.f (lane 9) extracts

Pan-Pbx antibody supershifts *prep1.1* /Pbx4 dimers present in nuclear of wild-type(lane 5) and *prep1.1* morphants 24 hpf zebrafish embryos (lane 8).

In the controls O-1 binds the dimers Prep1/Pbx1a and Prep1/Pbx1 b *in vitro* translated (lane 1), and binds the dimers present in nuclear extracts of HEK293 cells.

obtained from 24 h.p.f. WT and Prep1-morphants As a control, I used *in vitro* co-translated human Pbx1a and Pbx1b and Prep1. The pan-Pbx antibody identified the 50 and 40 kDa bands expected for the two Pbx isoforms when compared with the *in vitro* translated proteins (Fig. 6.6, lanes 1 and 2). Moreover, the same two bands were identified in extracts from human HEK cells, while only one of the two bands, a species of approximately 40 kDa, was revealed in an extract from mouse testis (Fig. 6.6, lanes 3 and 5; Wagner *et al.*, 2001). As showed in figure 6.6, the pan-Pbx antibody revealed two major bands in 24 h.p.f zebrafish embryos corresponding to Pbx2 and Pbx4 (which are both expressed at this stage of development); the intensity of each band was stronger in nuclear than in cytoplasmic extracts (compare lane 7 and 8 or 9 and 10). Moreover, the ratio between nuclear/cytoplasm Pbx proteins, of Pbx2 (Fig 6.6, lane 10) decreased significantly in the *prep1 .1* morphants. These data suggest that also in zebrafish, *prep1.1* might stabilize Pbx proteins and serve to increase the nuclear localization of different Pbx proteins. Furthermore, I tested by EMSA the presence of the functional eterdimers Prep1-Pbx4/*lazarus* in the nuclear extracts from wild type and *prep1.1* morphants at 24 h.p.f. As shown in figure 6.7 a band comigrating with Prep1/Pbx-b heterodimers binds to the O-1 oligonucleotide in WT (lane 6) nuclear extracts. On the contrary O-1 oligonucleotide weakly binds heterodimer in the MO nuclear extracts (lane9). Both in WT and MO the Prep1-Pbx4/*lazarus* dimers were specifically supershifted by pan-Pbx serum (lanes 8 and 11), suggesting the presence of a functional Prep/Pbx dimers. On contrary the presence of anti-Prep1 antibody do not have effect. Probably the antibody produced against the human Prep1 is not able to recognize zebrafish *prep1.1*(anti-prep1 antibody was unable to recognize *prep1.1* also in Westernblot experiment, data not shown).

Discussion Chapter 6

Prep1 knockout mice present embryonic lethality.

Prep1 gene was inactivated by a gene trapping approach, inserting a targeting vector in the *Prep1* locus (Fig 6.1), confirmed both by Southern blot and by PCR analysis with specific primers (not show). The insertion occurs in the first intron, before the ATG (Fig 6.2). The presence of a complete ATG could represent a problem, producing a normal Prep1 protein. However, the absence of Prep1 protein in the tissues of Prep1-/-mice confirms that this ATG does not work and indicates that the inactivation of Prep1 by gene trapping generated real knockout mice (Fig 6.3). Furthermore, all blotting data have carried out with nuclear extracts only; citoplasmatic extracts also need to be analyzed.

Prep1-/- mice obtained from the first intercrosses did not show any evident phenotype, but as I proceeded with the genetic background cleaning. I observed a consistent reduction in the number of Prep1-/- pups.

Mice lacking Prep1 function have severe morphological abnormalities (Fig 6.4) and preliminary results show that Prep1-/- embryos died *in utero* between E15.5 and E16.5 d.p.c. The strong phenotype shown in figure 6.4 was observed in 80% of Prep1-/- embryos. In 20% of Prep1-/- embryos the phenotype was less severe.

In order to understand the reason behind this variability I need to obtain a congenic background by backcrossing with C57Bl6/J wild type mice for 9-10 generations.

Similarity between Prep1 and Pbx1 knock out phenotypes.

It is important to note that the abnormalities observed in Prep1-/- embryos and those presented by Pbx1-/- embryos are similar. Genetic inactivation of Pbx1 resulted in severe hypoplasia in lungs, liver, stomach, gut, kidney and pancreas, aplasia in spleen, and in ectopic organs in. thymus and kidneys. Furthermore, Pbx1-/- embryos show defects in the axial and appendicular patterning, including morphological transformations of skeletal elements from the second branchial arch to the first. The absence of Pbx1 also causes abnormalities in chondrocyte proliferation and differentiation, resulting in precocious ossification and bone formation (Selleri *et al.*, 2002).

Recent data obtained in our laboratory clearly indicate a cross-requirement for stabilization and nuclear localization between Prep1 and Pbx2 proteins (Longobardi *et al.*, 2003, submitted). In accordance with these observations, the inactivation of *prep1.1* in zebrafish using the morpholin approach revealed that also in zebrafish, *prep1.1* might stabilize Pbx proteins and increase their nuclear localization. Indeed, the ratio between nuclear/cytoplasm Pbx proteins, in particular Pbx2, decreases significantly in the 24 hpf *prep1 .1* morphants (Fig 6.6, lane 10). In zebrafish embryos at 24 h.p.f. *prep1.1* and Pbx4/*lazarus* physically interact forming heterodimers binding the O-1 oligonucleotide, suggesting an in vivo functional role (Fig 6.7) Moreover, *prep1.1* morphants showed abnormal expression of several hindbrain markers (*mariposa, krox20, pax2.1, pax6.1*), including anteriorly expressed Hox proteins (*hoxb1a, hoxa2, hoxb2*). Finally, *prep1.1* morphants lacked all pharyngeal cartilages, an effect which was not due to the absence or impaired migration of neural crest cells into the pharyngeal arches but by the incapacity of these cells to differentiate into chondroblasts, thus indicating that *prep1.1* plays an essential role in hindbrain segmentation and craniofacial chondrogenesis.

All the abnormalities observed in the *prep1.1* morphants can be used as an additional model to direct the analysis of the Prep1-/- phenotype in mice. Preliminary results show that Pbx protein levels are significantly affected in Prep1-/- mice in a tissue dependent manner (Fig 6.5). These data suggest that Prep1 specifically regulates Pbx level in different tissues. Furthermore, in order to study the genetic interaction between Prep1 and Pbx1 I also plan to generate double heterozygous embryos.

All these observation could represent an interesting guideline to study the Prep1-/- mice. Therefore I plan to focus my future study on skeletal development, including the morphogeneis of mesenchymal structures derived from the branchial arches and limb buds.

Despite the similarity observed between Prep1, Prep2 and Meis proteins, which may suggest a possible redundancy in their functions, Prep1-/- mice show an abnormal phenotype, suggesting a specific and individual contribution of Prep1 in mouse development.

The data presented in this chapter are the results of preliminary studies and will be confirmed by additional experiments using mice from a congenic background.

The data presented in this chapter includes the preliminary experiments that I obtained with Prep1-/- mice and the work was carried out in collaboration with Francesco Argenton's group.
Zebrafish data are presented in the paper: De Florian, G., Tiso, N., Ferretti, E., Blasi, F., Bortolussi, M. and Argetoon, F. (2003, submitted). "*prep1.1* has unique and essential genetic function in hindbrain development and neural crest differentiations".

Chapter 7

FINAL DISCUSSION and FUTURE PERSPECTIVES

Members of the Hox gene family are sequence-specific DNA binding transcription factors which show high conservation within the DNA binding homeodomain. Consequently, the DNA binding properties of Hox proteins are very similar, raising the question of how proteins with equivalent biochemical properties can achieve *in vivo* distinct regulatory effects to define precise developmental programs (Ghering *et al.*, 1994; Lu *et al.*, 1996).

The formation of heterodimers between homeodomain proteins is an important and general mechanism for generating specificity *in vivo*. The interactions with cofactors belonging to the PBC family seem to be the key to the Hox binding specificity *in vivo*. Although very few *in vivo* target sites of Hox proteins are known, Hox auto-cross and para regulatory elements have been shown to be regulated by functional interactions with Hox-PBC dimers. Indeed PBC proteins form heterodimers with a specific subset of anterior Hox proteins (Chen *et al.*, 1997). Furthermore Meinox and PBC proteins form stable DNA –independent dimers (Berthelsen *et al.*, 1998). Member of the Meis family (Meis1) can also directly bind to some posterior Hox proteins (Shanmugam *et al.*, 1999; Shen *et al.*, 1999).

Therefore the interactions between PBC, Meinox and Hox proteins give rise to a different number of possible complexes, suggesting that combinatorial mechanisms are involved.

PBC and Meinox proteins are widely expressed, although individual family members may have a temporal and spatial pattern. The mechanism regulating the subcellular localization and stabilization of these proteins begins to be understood. The spatial and temporal restricted expression of *Hox genes* allow a specific restricted pattern of activity leaving open the question of how this segmental restricted expression pattern is established.

The Prep class of genes.

Prep1 is a homeodomain containing protein belonging to a small class of the atypical homeodomain proteins family termed the TALE family (Burling *et al.*, 1997). Prep1 shows homology to members of the Meinox family, which include murine Meis1, Meis 2 and Meis 3, *C. Elegans* unc 62, *Drosophila* homotorax (Hth) and other orthologues in Zebrafish and *Xenopus Laevis* (Fig 1.18; Fig 3.2; Burglin *et a.*, 1997; Moskow *et al.*, 1995; Nakamura *et al.*, 1996; Oulad-Abdelghani *et al.*, 1997; Steelman *et al.*, 1997; Pai *et al.*, 1998). Sequence similarity is restricted to the level homeodomains and two protein-protein interactions domains located at the N-terminus and termed HR or MD domains (Fig 3.2).

Recently, a new member of the Meinox family was identified and named: Prep2. The generation a phylogenetic tree showed that PREP1 and PREP2 define a sub-family of Meinox proteins distinct from Meis and Hth (Fig 3.14 C and Fognani *et al.*, 2002). Protein sequence comparison between Prep1 and Prep2

reveals high homology between their functional domains (HD and HR; Fig 3.14A), however the similarity in all other regions highlights the existence of a subfamily different from that of Meis. However, one important functional difference between the two Prep proteins was demonstrated by transfection experiments, in which the transcriptional activity of the *b1ARE* regulatory region revealed a positive response to Prep1 transfection and a mild repressive effect to Prep2 transfection (Fognani *et al.*, 2002; Haller *et al.*, 2002) suggesting that these two proteins may have opposite functions.

The Meinox homeodomain is highly conserved during evolution and is present both in the animal and in the plant kingdom. However, while in the animal kingdom the Meinox-type homeodomain is rare, in plants, it is much more widespread. In the plant kingdom the Meinox –like proteins belong to the large Knotted class and include proteins that are involved in meristeme development (Deshpande *et al.*, 1999).

Expression of Prep1

In this thesis I present data showing that Prep1 is widely expressed in adult murine tissues and in embryos throughout embryonic development (Fig 3.6 and Fig 3.7). Moreover, DNA-binding Prep1-Pbx complexes were found in nuclear extracts purified from embryos and from ES cells. The early expression of Prep1 might be associated with the regulation of genes involved in dynamic cellular processes occurring in early development. Both in embryos and in adult tissues Prep1 represents a functional partner of Pbx protein, binding all the Pbx isoforms (Fig 3.4 and Fig 3.8). Prep1-Pbx complexes were present in all tissues tested. In addition, EMSA experiments revealed that Prep1-Pbx dimers present a tissue-specific subunit composition (Fig 3.4 and Fig 3.5). Thus, while Prep1 is ubiquitously present, its tissue-specificity may depend on Pbx partners, suggesting differential activities for the different complexes. However, the exact role of different Prep1-Pbx dimers is not understood.

Immunohistochemistry experiments performed on sections of adult murine organs revealed that Prep1 is localized in the nuclei of most cells (Fig 3.9, Fig 3.10 and Fig 3.11). There are considerable evidences supporting the idea that Prep-Pbx dimers must be localized in the nucleus in order to be active. Interestingly, oocytes have both Prep1 and Pbx proteins in the cytoplasmic compartment. This is probably indicative of the inactive transcriptional status of these cells (Fig 3.11).

In addition, a similar behavior is observed for *prep1.1* in zebrafish oocytes where *prep1.1* is present as maternal messenger (De Florian *et al.*, submitted). We could hypothesize that Prep/Pbx translocation into the nucleus occurs after the fertilization at time when active transcription starts. Prep1 is broadly expressed during murine embryonic development and localizes to the nuclear compartment of the majority of the cells showing strong correlation with gene expression and nuclear localization of Pbx proteins. I observed that the Pbx forms present in the Prep-Pbx dimers change throughout development (Fig 3.8). Unlike Prep1, different members of the PBC family have differential spatial and temporal expression, suggesting that Prep1 affects the function of multiple Pbx proteins (Fig 3.8A).

Thus, since Prep1 and Pbx proteins are co-expressed in the embryo at the moment when Pbx is also modulating Hox expression, we can hypothesize that Prep1 can regulate PBC-Hox activity. Indeed, immunoistochemistry experiments show that Prep1, Pbx1, Pbx2 and Hoxb1 proteins co-localize in r4 of the developing murine hindbrain (Fig 4.1). In addition the presence of Prep-Pbx in organs where the Hox proteins are not expressed suggest that these dimers could have additional functions The identification of target sites that bind Prep-Pbx complexes, as the *urokinase* enhancer, in the absence of Hox, supports this hypothesis.

What is the function of the Prep-Pbx–Hox ternary complex?

Dissection of the Hox regulatory elements both in Drosophila and mouse has provided important insights into Hox protein function. A bipartite consensus "TGATNNAT" was identified as a Hox/PBC consensus binding site, PH, able to regulate the segmental expression of Hox proteins. In collaboration with the Krumlauf. laboratory, we identified a *cis* PBC/Meinox element (PM) near a bipartite PH site in *Hoxb1*, *Hoxb2* and *Hoxa3* enhancers and we showed that those sites represent a combined conserved motif. I demonstrated that Prep1 increases the binding selectivity of Pbx-Hoxb1 complexes to these sites. Furthermore, the cooperative interactions between Prep1, Pbx and Hox proteins are important in orchestrating the binding of multimeric complexes to the combined PM-PH sites both *in vitro* and in cell culture. *In vivo* it showed to be relevant and essential for r4-restricted expression of *Hoxb2* in the developing hindbrain and may also contribute to aspects of *Hoxb1* expression in r4 and *Hoxa3* in r5/r6 (Popperl *et al.*, 1995, Gould *et al.*, 1997, Maconochie *et al.*, 1997, Li et al., 1997, Pinsonneault *et al.*, 1997, Jacobs *et al.*, 1999; Ferretti *et al.*, 2000, Manzanares *et al.*, 2001).

The work proposed in chapter 4 highlights the key role played by the interactions of Prep1, Pbx and Hox proteins with combined PM-PH target sites in determining the spatially restricted pattern of expression of *Hox genes*. We propose that a Prep1-Pbx-Hoxb1 ternary complex bound to the combined PM-PH sequence acts as an activator of *Hoxb2* restricted expression in the hindbrain. We have also showed that this complex has a functional role *in vivo*, directing the expression of *Hoxb2* in r4 (see chapter 4; Ferretti *et al.*, 2000). This is the first evidence describing a functional role for the ternary complex *in vivo*.

Although other types of ternary complexes have been observed, the Prep1-Pbx Hox ternary complex shows peculiar characteristics. In contrast to other ternary complexes described (Fig 1.19), the formation of the Hoxb1-Pbx1-Prep1 complex requires all the DNA-binding domains of the participating proteins. In addition, the ternary complex formation requires two different DNA-binding sites, PM and PH (Fig 4.3 and 4.4). A similar situation is described in *Drosophila*, where a Labial-Exd-Hth ternary complex activates the *labial* promoter, requiring DNA-binding of all three proteins (Ryoo *et al.*, 1999). Thus, this kind of DNA-binding ternary complex seems to activate the transcription of specific genes, and it is also conserved during the evolution of the fly to mammal. I also observed that the distance between, PM and PH sites, in different enhancers are extremely variable: 8 bp in *Hoxb2*, 13bp in *labial*, 14 bp in *Hoxa3* and 17 bp in *Hoxb1*. In addition, as with in labial, the position of these sites is variable (Fig 4.19).

All these observations suggest that there may be significant flexibility in the assembly of the three proteins on these elements. This flexibility may be important for the recruitment of additional factors, which may be different for different enhancers and thus could confer further level of enhancer-specificity to the ternary complex.

Like Prep1, Meis can also form a ternary complex and bind the PM-PH sequence (Jacobs *et al.*, 1999), suggesting that both Prep1-Pbx-Hox and Meis-Pbx-Hox can co-exist. The mechanisms that modulate the binding and the activity of these two complexes are as yet not fully understood.

How do Prep and Pbx participate in the segmentally restricted expression pattern of *Hox genes* in the hindbrain?

How the ubiquitous factors Prep and Pbx can direct the segmentally restricted expression pattern in the hindbrain remains to be clarified. The fact that Prep1 changes its Pbx partners in different tissues and during embryonic development could confer a high degree of specificity of the dimers (Fig 3.8). Furthermore different Prep1-Pbx dimers may have a different affinity for PM and PH sites or could also act as both activators and inhibitors. In addition we need to consider that Prep1 and Meis are probably competing for binding to the Pbx proteins and that the two dimers Prep-Pbx or Meis-Pbx likely different functions *in vivo*. Indeed, Meis is able to interact directly with posterior Hox. Finally, different components of dimers may recruit enzymatic activities (HAT v HDAC) and/or different co-activators or corepressors.

Taken together these observations may explain how ubiquitous factors, such as Prep and Pbx, could modulate *Hox genes* expression in a spatially restricted manner. However, *Hox genes* expression is known to be spatially restricted by the activity of morphogens and additional transcriptional factors (Krumlauf 1994; Mann *et al.*, 1998; Manzanares *et al.*, 2001). Therefore, the presence of expression-restricted Hox proteins may in turn specify the activity of the Prep1-Pbx-Hox complex. Thus, between the Hox and TALE proteins which are factors and which the cofactors?

What is the role and DNA-binding specificity of the PM site?

Although we observed the formation of Prep1-Pbx-Hox complexes on the combined PM-PH motif in different enhancers, we also observed that the behavior of these binding sites *in vivo* is different. We demonstrated an essential *in vivo* requirement for the PM site only in the case of the *Hoxb2* and not *Hoxb1* or *Hoxa3* enhancer (Fig 4,14 and Fig 4.15). These results do not exclude a role for the PM site in expression of Hoxb1 in r4 or in expression of *Hoxa3* in the r5/r6 in murine hindbrain. Comparing the structure of these three enhancers, we found that while *Hoxb2* has only one PH site, *Hoxb1* and *Hoxa3* have three and two PH sites, respectively. Thus, *Hox* enhancers vary both in number and arrangement of their PH and PM sites. Additional copies of either of these two sites might act in a redundant manner to ensure and reinforce levels, in spatially restricted *Hox genes*, or to modulate the ternary complex function. By analyzing of different PM sites, we conclude that they have different DNA-binding specificity and affinity for the Prep1-Pbx complexes. Moreover, Prep1-Pbx complexes can also bind to

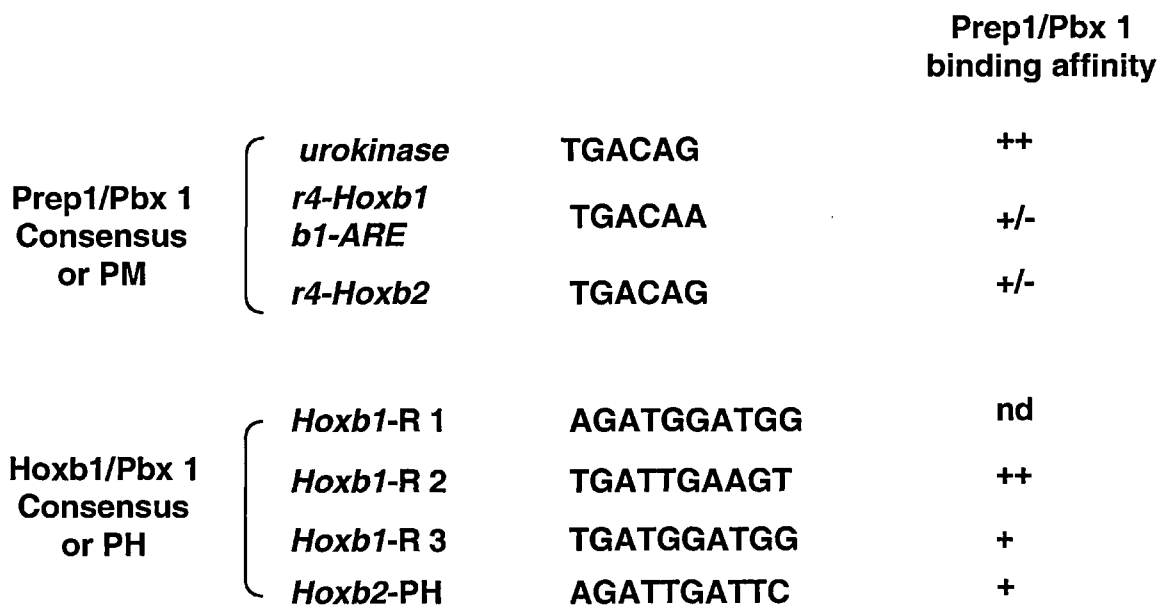


Figure 7.1
The Pbx-Meinox binding sites (PM) of *urokinase*, *Hoxb1* and *Hoxb2* enhancers show a different binding specificity for Prep1-Pbx complexes. Only the PM site in the *urokinase* enhancer bind Prep1-Pbx dimers with high affinity.
In addition, the Pbx-Hox binding sites (PH) of *Hoxb1* and *Hoxb2* enhancers also bind the Prep1-Pbx dimer.
The comparison of PM and PH binding sites suggest that the new consensus for the binding of Prep1-Pbx dimers is T/AGANNG.

the PH site, originally described as a Pbx-Hox binding motif. For this reason I compared the different PM and PH sites to define a new consensus motif for the binding of Prep1-Pbx dimers, namely "T/AGANNG". I also demonstrated that the 3 bases located upstream to the T/AGANNG consensus sequence play an essential role in determining the affinity of PM for Prep1-Pbx. The presence of CTC confers high affinity to the consensus sequence as in the case of the *urokinase* enhancer (Fig 4.11).

I noticed that the PM site (R2) in the *Hoxb1* enhancer, which can inhibit the formation of the ternary complex, has the CTC motive upstream of its core sequence. In fact, R2 is a strong Prep1-Pbx binding site (Fig 7.1). On the other hand, I also noticed that both in the *Hoxb1* and *Hoxb2* enhancers the PM site belonging to the PM-PH motif is always a low affinity PM site. Is this condition a necessary requirement for the formation of the ternary complex?

The data I present in chapter 4 (Fig 4.12) show that by converting those PM sites into high affinity PM site inhibits the ternary complex formation, while enhancing the binding of Prep1-Pbx dimer. Thus, it is possible that low affinity PM sites are entry point to built a ternary complex in collaboration with a PH site, while high affinity binding of Prep1-Pbx dimer would prevent this possibly.

Different roles of PH sites in Hoxb1 regulation.

I have demonstrated that the PM-PH motif of the *Hoxb1* and *Hoxb2* enhancers promote the *in vitro* formation of the ternary complex and *in vivo* are instructive sequences capable of driving the expression of reporter gene in a r4-restricted way (Fig 5.13 and 5.14). The PM-PH motifs in *Hoxb1* and *Hoxb2* are substantially different in terms of sequences, spacing between the two sites and binding of additional proteins, but they have the same behavior *in vivo*. If the PM-PH sequences are sufficient to drive the expression of Hoxb1 in r4 what is the role of the additional PH sites? And why is the PM site not required in Hoxb1 enhancer activity.

According to the above observations we can no longer consider the multiple PH sites as redundant sites as binding properties. R3 is able to bind Pbx-Hoxb1 as well as Prep1/Pbx dimers, albeit with low affinity. In addition, in cooperation with PM, R3 is involved in the formation of the ternary complex. On the contrary, R2 acts as a high affinity binding site for Prep1-Pbx (like a high affinity PM site) that exerts its inhibitory effect both *in vitro* on the ternary complex formation (Fig 5.5) and *in vivo* on r4-restricted *Hoxb1* expression (Fig 5.15). R1 activity needs further investigation

The multiple PH sites may also be involved in the formation of additional complexes with other transcription factors like Oct1, belonging to the POU family. Indeed, in addition to the three PH sites, I showed that Oct1 binds an octamer-like sequence located between R2 and R3 (Fig 5.5). Several data suggest that the POU domains can modulate protein-protein interactions resulting in the formation of homodimers and complexes with other transcriptional regulators (Ryan and Rosenfeld, 1997). Therefore POU transcription factors interact with Prep1, Pbx and/or Hox proteins. As in the *Hoxb1* enhancer, the presence of both PM and Oct1 binding sites was also observed in the *urokinase* (Fig 5.19), suggesting a conservation off an interaction mechanism involving Oct1.

Effects of Prep1 inactivation.

Both in zebrafish and mouse Prep1 has an essential role in embryonic development, although the data in mice are partial and preliminary.

Inactivation of *prep1.1* in zebrafish using the morpholino approach results in severe alterations of the segmentation pattern of hindbrain due to the modification of anteriorly expressed Hox proteins (*hoxb1a, hoxa2, hoxb2*). Furthermore, *prep1.1* morphants lacked all pharyngeal cartilages. This phenotype depends on the incapacity of neural crest cells to differentiate into chondroblasts, indicating that *prep1.1* plays an essential role in hindbrain segmentation and craniofacial chondrogenesis (De Florian *et al.*, 2003 submitted).

Prep1-/- mice showed a lethal embryonic phenotype, dieing *in utero* at E 15.5-E 16.5 d.p.c. Gross morphology analysis of Prep1-/- embryos revealed massive subcutaneous edema, generalized pallor diminished vascularization, smaller livers and abnormal orientated forelimbs (Fig 6.4). So far, the abnormalities observed in Prep1-/- embryos and those presented by Pbx1-/- embryos are similar, suggesting a genetic interaction between these two factors Much evidence points to a strong correlation between the levels of Meis and Pbx proteins, but all the evidence is accumulating with use of dominant-negative, hence ortholog non specific, mutants. I demonstrated that both in zebrafish and in mouse the inactivation of Prep1 leads to the down regulation of nuclear Pbx, consistent with a role for Prep1 in controlling Pbx stability. In addition, analysis of tissues of Prep1-/- mice revealed that the down regulation of Pbx proteins occurred in a tissue-dependent manner, different Pbx proteins being down regulated in different tissues (Fig 6.5).

The fact that Prep1 is involved in the protein stabilization of all three Pbx forms could suggest a stronger phenotype for Prep1-/- mice. In order to further investigate the genetic interaction between Prep1 and Pbx1 we are also planning to generate the Prep1/Pbx1 heterozygous embryos.

The phenotype observed in the morphants *prep1.1* and that observed in the Pbx1-/- mice may help to direct the analysis of the Prep1-/- phenotype. We are planning to analyze the head skeletal development and the appendicular skeletal formation, focusing on the morphogenesis of mesenchymal structures derived from the branchial arches and limb buds. Furthermore, we plan to examine the perturbation of the segmentation pattern of hindbrain by analyzing the Hox expression patterns in the Prep1-/- embryos. However, we must also consider the possibility that Prep1 may plays Hox-independent roles in tissues and functions not expressing Hox.

In addition, because the similarity observed between members of the Meinox family, which suggests a possible redundancy in their functions, Prep1-/- mice may show specific phenotypes which need to be identified. These data strongly suggest that Prep1 may have a unique and individual role in mouse embryonic development. In order to further clarify the role played by Prep genes family in embryonic development we are planning to also knockout the Prep2 murine gene.

The analysis of the phenotype of the single Prep1-/- knockout mouse as well as the generation and the study of the double Prep1/Prep2 knock-out, will define the role of Prep proteins in embryonic development. In addition, the comparison with the Meis knockout mice phenotypes when available will give

us important information about the specific function played by Prep among the Meinox proteins. The data presented in this thesis are the results of the preliminary analysis of the mouse Prep1-/- phenotype and need to be confirmed by additional experiments.

In summary, the data collected so far show that Prep1 has functions.

1. Prep1 prevents nuclear export of Pbx proteins (Berthelsen *et al* 1999)

2. Prep1 modulated the amount of Pbx protein by stabilizing the Prep-Pbx heterodimer (De Florian *et al.*, 2003, submitted and Longobardi *et al.*, 2003, submitted).

3. Prep1 participates in the formation of transcriptionally active complexes, which modulate the activity of Pbx-Hox dimers and promote Hox expression in murine hindbrain (Ferretti *et al.*, 2000).

4. Prep1 may act as a cofactors of chromatin modifiers, such as Oct1 and Sox, and thereby modulate Hox expression and Pbx-Hox activity (Ferretti *et al.*, manuscript in preparation).

5. Prep1 is a co-regulator of Pbx1 protein in skeletal development (Ferretti *et al.*, in progress)

Chapter 8

Materials and methods

8.1 Recombinant DNA Techniques

8.1.1 Reagents and standard method procedures.

All basic DNA standard methods, extraction, purification, ligation, labeling, bacterial cultures and transformation, media and buffer preparation were carried out according to methods described by Sambrook *et al.*, 1989. Restriction enzymes and DNA-modification enzymes were obtained from Boeringher Mannheim and Promega and used according to manufacture's instructions.

8.1.2 Competent bacterial cells.

Competent cells were made from *E.Coli* strain:
-CN524 (Peterson, 1989), with the F factor from strain 7118 and the genotype ara D139D(araABOIC-leu)7679 galU KD(lac)X74 rspL thi recA1/F' proAB+ lac IqlacZDM15.
-HB101 with the genotype supE44 hsd (rB-mB-) reaA13 ara14 proA2 lacY1 galK2 rpsL20 xyl5 mtl1 (Bolivar and Backman, 1979).
-DH5α with the factor f80 lacZDM15 that permits α-complementation with the amino terminus of β-galactosidase and has the genotype supE44 Dlacu169 (f80lacZDM15) hsdR17 recA1 endA1 gyrA96 thi1 relA1.
Cells were made competent by the $CaCl_2$ method.
Transformation was performed by incubation of bacteria for 30 min on ice, heat-shock for 90 seconds at 42°C and plating on LB-broth agar containing the appropriate antibiotic (ampicillin or kanamycin).

8.1.3 Plasmid isolation

Cells were grown at 37°C with overnight shaking.
Small-scale plasmid preparations "mini-preps" (3ml culture) were made by the alkaline lysis method (Sambrook *et al.*, 1989). Large preparation "maxi-preps", were purified on Qiagen columns (Qiagen, USA). In this case bacteria were grown overnight in 500 ml of LB medium plasmid DNA was extracted the following day. DNA preparation and concentration were determined both by restriction enzyme digestion and spectrophotometer measurement at 260 nm.

8.1.4 Sequencing

All plasmid sequences were controlled by automated sequencing carried out by PRIMM, Italy.

8.1.5 Cloning.

Restriction fragments were separated by gel electrophoresis in 0.5xTBE buffer (0.045 M Tri-borate, 0.01 M EDTA). DNA fragments were purified from the agarose using the Quia-quik DNA purification system. DNA manipulations and storage were in TE buffer 10 mM Tris, 1 mM EDTA).

8.1.6 PCR

PCR reactions (Polymerase Chain Reaction) were performed using Taq polymerase and buffers supplied by the manufacturer (Boeringher Mannheim). Standard PCR reactions were carried out for screening purposes, including cloning verification, colony screening (Zon *et al.*, 1989) and genotyping screening. All reactions were performed using 1X PCR reaction buffer, 1μM of each primer, 0.2μM deoxynucleotides (dNTPs), and 1.5u Taq polymerase in total volume of 50 μl.

PCR colony screening reactions were generally performed using the following parameters: an initial denaturation (5min 95°C) followed by 30 cycles consisting of denaturation (30 seconds, 95°C), annealing (30 seconds, 56°C) and elongation (2 minutes, 72°C), and a final extension step (10 minutes 72°C). Individual fresh bacteria colonies were resuspended in 25 μl of PCR buffer containing the appropriate primers and transferred on LB-plate in order to maintain a replicate colonies. Positive samples were identified by gel electrophoresis and colonies corresponding to positive samples were inoculated for mini-prep plasmid isolation.

-Preparative PCR reactions for expression vectors and mutants were performed using a high concentration of template (3μg) to minimize the introduction of mutations, and low numbers of cycles (typically 12-15) were performed using "Expand High Fidelity" polymerase (Boeringher Mannheim).

8.1.7 RNAase Protection Assay

For RNAase protection analysis, a fragment of mouse Prep1 cDNA (from position 1052 to 1263) was cloned into pBluescript (Stratagene). Prep1 antisense RNA was produced using ribo-probe Kit (Promega) in the presence of 32P-GTP (Amersham). Labeled transcripts were purified by acrylamide gel electrophoresis. RNA extraction and RNAase protection assay was performed following standard methods (Sambrook *et al.*, 1989).

8.1.8 Genomic DNA purification from cell culture and mice-tails

A small piece of tail was placed in 750 μl of digestion buffer (50mM Tris-HCl pH8, 100 mM EDTA, 100 mM NaCl, 1% SDS, 0.5 mg/ml Proteinase K, and incubated overnight at 55°C. The day after, RNAase was added to the reaction and incubated for 1 hour at 37°C. The DNA was purified from proteins by several extractions with phenol/chloroform and precipitated adding 500μl of ethanol (100%) to 500 μl of clear supernatant. After centrifugation, the pellet was washed with 70% ethanol and allowed to air dry. The pellet was resuspended in 100 μl TE 1X. Genomic DNA was stored at 4°C. This procedure generally yielded 50-100 μg of genomic DNA.

8.1.9 Southern analysis

Genomic DNA was incubated with appropriate restrictions enzymes and run on a 0.8% (w/v) agarose gel in 1X TAE buffer. The gel was denaturated for 45 min. in 0.5 M NaOH, 1.5 M NaCl and neutralized 45 min. in 1 M Tris pH 8, 1.5 M NaCl. DNA was transferred overnight onto a nylon positively charged membrane (Amersham) by capillary action in SSC 10X.

The day after the DNA fragments were fixed on the membrane using UV-cross linker. The filter was incubated with a P^{32} labeled probe at 65°C in Hybridization-plus buffer (Sigma). The day after the filter was washed three times by increasing the stringency to remove non-specific signal (from 2X SSC, 0.1% SDS to 0.1X SSC 0.1% SDS). The membrane was finally exposed to Kodak film overnight at –80°C.

8.1.10 Radiolabeling of DNA for Southern and Northern Blotting.

The probe used for Southern and Northern blotting experiments was prepared by PCR, using the murine Prep1 cDNA as target. The primers employed for PCR were:
-probe 2-3 forward primer (2) 5'ATGATGGCGACACAGACGCTAAGTATA3' and reverse primer (3) 5'GGGGTCTGAGACTCGATGGGAGGAGGACTC3'.
-probe 8-9 forward primer (8) 5'GCAACACCCAGCGAGCTTGCCATG3' and reverse primer (9) 5'GGAGTCGCTGTTCTCCAGCAGCACCAG3'.
To obtain a probe containing the entire coding region of murine Prep1, (probe 2-9) a combinations of primers 2 and 9 was used.
100 ng of the DNA probe were labeled using the High-Prime Amersham labeling kit, based on the random primed DNA-labeling method, using [α-^{32}P] dCTP according to manufactures instructions. The radio-labeled DNA fragments were purified by passing the labeling reaction mixture through a packed Sepharose G-50 column (Boehringher Mannheim).

8.2 Protein Techniques

8.2.1 Immunohistochemistry

Immunohistochemistry on 6 μm paraffin sections was performed using anti-Prep1 (dilution 1:5,000), anti-Pbxs (dilution 1:200), anti-Hoxb1 (dilution 1:500) and PI (pre-immune serum as a negative control; 1:5,000 dilution) with the Vectastain peroxidase staining kit (Vector Laboratories, USA) according to manufacturer's instructions. For epitope unmasking, sections were treated for 10 min at 37°C with trypsin and subsequently blocked in blocking solution (90% of fetal calf serum: FCS and 1% BSA). Incubation with the primary and secondary antibodies was performed in 10% FCS and 1% BSA.

8.2.2 Nuclear extract preparation

Tissue extracts
For EMSA and immunoblotting assays nuclei were prepared from adult murine (CD1) tissues and from dissected mouse embryos at the indicated stages.
After dissection, tissues were washed with PBS 1X to remove all remaining blood and frozen in liquid nitrogen. The tissues pieces were crushed into a powder using a mortar and pestel. The resulting powder was resuspended in lysis buffer: buffer A (10 mM HEPES pH 7.9, 30 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$), kept on ice for 10 min and supplemented with addition of Triton X-100 to a final concentration of 0.1%, to lysate the cells. The nuclei and the cells debris were collected by centrifugation. The supernatant was removed to a new tube, 0.11 vol of buffer B (0.3 mM HEPES pH 7.9, 1.5 mM MgCl2) were added and the samples were centrifuged. The resulting supernatant corresponded to the cytoplasmic extract (CE). Nuclear extract (NE) was prepared

by resuspending the nuclear pellet in 60 µl buffer C (20 mM HEPES pH 7.9, 25% glycerol (v/v), 0.42 M NaCl, 1.5 mM $MgCl_2$, 1 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$) and incubation at 4°C with shaking for 30 min. The extracts were cleared by centrifugation. Finally, the extracts were aliquoted and stored in liquid nitrogen.

Cells extracts

Cells were grown to confluence in 10 cm dishes under standard condition (see below), washed with PBS and collected in a tube. Nuclear and cytoplasmic extracts were prepared as described above for tissues extracts.

8.2.3 Protein concentration determination

Protein concentration was measured using the Bradford Assay according to manufacturer's instructions (BioRad protein Assay). An appropriate dilution of gamma globulin was used as a standard(BioRad).
Protein concentration was quantified by measuring the absorbance at 595 nm.

8.2.4 *In vitro* transcription/translation

All cDNAs used were cloned into the pSG5 plasmid and were under the control of the T7 promoter. The cDNA were transcribed and translated by using the TNT coupled transcription/translation reticulocyte lysate kit (Promega) according to the manufacturer's instructions using T7 polymerase and ^{35}S-methionine (Amersham). PREP1 and PBX were co-translated by using plasmids in equimolar amounts.
Translation products were analyzed by SDS-PAGE and visualized by exposure to Kodak X-OMAT film.

8.2.5 SDS-PAGE

Protein extracts were separated by SDS-PAGE (0.75 mm thick) cast and run in a Hoefer Mighty Small SE250 gel apparatus at 20mA/gel until the bromophenol blue front was about to exit the gel.
Resolving gels were prepared from an acrylamide stock solution (30% acrylamide, 0.8% bis-acrylamide) and separation buffer (4X stock solution: 1.5 M Tris-HCl pH 8.8, 10% SDS, Sambrook *et al.*, 1989). For the stacking gel 5% acrylamide was used and a stacking solution with Tris-HCl pH 6.8.
The running buffer used f or electrophoresis contained 0.025 M Tris, 0.192 M glycine and 0.1% SDS, pH8.3.
Samples were loaded in an equal volume of 2X sample buffer (20 mM Tris-HCl pH 8, 4% SDS, 200 mM DTT a.1% bromo phenol blue(BPB) ON heated at 100°C for 5 minutes, than centrifugated briefly, cooled to RT and loaded on the gel.

8.2.6 Antibodies

Antibodies against PBX proteins were obtained from Santa Cruz Biotechnology (Santa Cruz, CA). The antibodies anti-Pbx1, anti Pbx2 and anti-Pbx3 recognize a peptide from the N-terminus of the proteins. The anti-Pbx1 and anti-Pbx3 recognize both splice variants. The anti-Pbx-a antibody recognizes a common C-terminal peptide in all 50 kDa splice variant of the PBX proteins. All PBX antibodies are produced against the peptide of the human protein, but are also able to recognize both human and mice PBX proteins.

The anti-PREP1 antibodies were produced using a GST-PREP1 fusion protein to immunize rabbit and purified on Protein A sepharose (Berthelsen *et al.*, 1998a) Anti-HOXB1 antibodies were produced by Babco (USA). The antibody was raised against a GST-HOXB1 fusion protein and purified by affinity chromatography. Anti Meis1 antiserum was a kind gift from A.M. Buchberg.

8.2.7Immunoblotting

After electrophoresis gels were blotted onto Immobilon-P polyvinyl difluoride (PVDF) membrane (Millipore) using a JKA semidry blotter with blotting buffer (25 mM Tris, 192 mM glycine, 10% methanol). Blotting was performed at 1 mA/cm^2 for 90 min. To determine the transfer efficiency, the proteins transferred onto the PDVF membrane were stained with Ponceau (Sambrook *et al.*, 1989). After the staining the filter was washed in (TBS-T: 150 mM NaCl, 50 mM Tris-HCl pH 7.5 and 0.1% Tween-20) and then incubated overnight at 4*C with blocking solution (TBS-T: 150mM NaCl, 50mM Tris-HCl pH 7.5, 0.1% Tween-20 and 5% non-fat dry milk).

Membranes were incubated with primary antibodies for 1hour at room temperature in blocking buffer (TBS-T: 150 mM NaCl, 50 mM Tris-HCl pH 7.5, 0.1% Tween-20 and 1% non-fat dry milk). Membranes were washed 5 times (2X1 min and 3 X 15 min) in TBS-T and probed for bound antibodies by incubation with horseradish peroxidase conjugated secondary antibodies (Amersham) diluted 1:5,000 in blocking buffer for 1 hour. Immunoblotting experiments were carries out with anti-PREP1 (1:8000 dilution), anti-PBXs (1:2000 dilution), anti-Meis1 (1:8000) dilution and anti-HOXB1 (1:1000 dilution) antibodies. The incubation with secondary antibody was carried out as details above.

The detection was done using the chemoluminescent substrate Kit (Super Signal, Pierce Chemical) according to supplier's manual.

8.2.8. EMSA

For the Electrophoretic Mobility Shift assays (EMSA) were performed using 10 µg of nuclear extract in a buffer containing 2 µg poly-dIdC in 9 µl H2K150 (25 mM HEPES pH 7.9, 20% glycerol (v/v), 150 mM KCl, 1 mM $MgCl_2$, 0.2 Mm EDTA, 1 mM DTT, 0.5 mM PMSF). This was incubated on ice for 10 min. After the addition of 30,000 c.p.m. ^{32}P-labelled oligonucleotide, and antibody when required, the reactions were left at room temperature for 10 min and analyzed by 5% PAGE in 0.5 X TBE.

EMSAs with *in vitro* translated proteins were performed using 2 µl of reticulocyte lysate containing the desired combination of proteins mixed with 18 µl of PPH binding buffer (10 mM Tris-Cl pH 7.5, 75 mM NaCl, 1 mM EDTA, 6% glycerol, 3 mM spermidine, 1 mM DTT, 0.5 mM PMSF, 1 µg poly (dIdC), 30,000 c.p.m ^{32}P labeled oligonucleotide to a total volume of 20 µl, when necessary antibody or unlabeled probes were added. After 30 minutes of incubation on ice the reactions were separated by 5% PAGE in 0.5 X TBE.

For the competition assays 50, 100 or 500-molar excess of unlabeled competitor oligonucleotide was added to the binding reaction 10 min before the labeled probe.

For analysis of dissociation rate, a 100-fold excess of unlabeled competitor oligonucleotide was added after 10 minutes of incubation with the labeled probe

and samples from the same reactions were taken at different times and immediately loaded on native gels.

8.2.9 Oligonucleotides

Oligonucleotides used in EMSA were obtained commercially from Biosense (Belgium) or PRIMM (Italy). Oligonucleotides were resuspended and diluted in TE. Double stranded oligonucleotides were prepared by mixing equal molar ratios of single strand oligonucleotides. The mixture was heated at 90*C for 5 min and slowly cooled at RT. Purification of double-strand oligonucleotides were performed using preparative acryl amide 15% in 0.5X TBE. The oligonucleotides were eluted from the gel in 500 µl TEN (10 mM Tris, 1 µM EDTA, 300 mM NaCl, pH8) with shaking at 37°C overnight. The oligonucleotides were recovered by ethanol precipitation. The pellet was washed, dried and resuspended in 50 µl TE.

8.2.10 Labeling of the 5′ end of oligonucleotides

0.5 pmoles of double stranded oligonucleotides were incubated with 1 µl of 10X PNK buffer (Boeringher Mannheim), 5 Units Polynucleotide Kinase (PNK, Boeringher Mannheim) and [γ-^{32}P]ATP 3000 Ci/mMole (Amersham). The reaction were incubated at 37°C for 40 min and then diluted to 50 µl with TE and purified on a packed Sepharose G-25 column (Boeringher Mannheim). The labeling efficiency was determined by counting 5 µl in a beta-counter (Beckman) by Cerenkov counting.

8.2.11 Sequences of oligonucleotides used for EMSA.

O1	5′-CACCTGAGAGTGACAGAAGGAAGGCAGGGAG-3′
O1m	5′-CACCTGAGAtTcACAGAAGGAAGGCAGGGAG-3′
o-SP1	5′-GATCGATCGGGGCGGGGCGATC3-′
b2-PM-PH	5′-GGAGCTGTCAGGGGGCTAAGATTGATCGCCTCA-3′
b2-PM	5′-ATCGGAGGGGGAGCTGTCAGGGGGC-3′
b2-PH	5′-GGGGCTAAGATTGATCGCCTC-3′
b2-M1	5′-GGAGCTcTtAGGGGGCTAAGATTGATCGCCTCA-3′
b2-M2	5′-GGAGCTGTCAGGGGGCTAAcgTTcgTCGCCTCA-3′
b2-M1-M2	5′GGAGCTcTtAGGGGGCTAAcgTTcgTCGCCTCA3′
b1-PM-R3	5′-TCTTTGTCATGCTAATGATTGGGGGGTGATGGATGGGCGCTG-3′
b1-PH	5′-GGGGGTGATGGATGGGCGCTG-3′
b1-PM	5′-AGTGTCTTTGTCATGCTAATGATTGGG-3′
a3-PH1	5′-GCCGAGTCATAAATCTTGCCCAG-3′
a3-PHP1	5′-GCCGAGTCATAAATCTTGCCCAGCCATAAATGACAAAAACCATT-3′
a3-PP2	5′-GTGCGGGTTGATTATTGACCCACGCCTTC-3′
R2-PM-R3	5′-TCAGAGTGATTGAAGTGTCTTTGTCATGCTAATGATTGGGGGGTGATGGATGGGCG-3′
R2-pm-R3	5′-TCAGAGTGATTGAAGTGTCTTTcTtATGCTAATGATTGGGGGGTGATGGATGGGCG-3′
r2-PM-R3	5′-TCAGAGTcgTTcgAGTGTCTTTGTCATGCTAATGATTGGGGGGTGATGGATGGGCG-3′
R2-PM-r3	5′-TCAGAGTGATTGAAGTGTCTTTGTCATGCTAATGATTGGGGGGTcgTGcgTGGGCG-3′
R2-PM	5′-TCAGAGTGATTGAAGTGTCTTTGTCATGCTA-3′
Flop1 (PM-R2)	5′-TCTTTGTCATGCTAATGATTGGGGGGTGATtGAaGGGCGCTG-3′
Flop2 (R3-PM)	5′-TCAGAGTGATgGAtGTGTCTTTGTCATGCTA-3′
Flip1 (R3-PM-R2)	5′-TCAGAGTGATGGATGTGTCTTTGTCATGCTAATGATTGGGGGGTGATTGAAGGGCG-3′
Flip2 (r3-PM-R2)	5′-TCAGAGTcgTGcgTGTGTCTTTGTCATGCTAATGATTGGGGGGTGATTGAAGGGCG-3′

8.3 Cell culture methods

[1]All tissue culture media, equipment and chemicals were supplied by Gibco-Life Technology, USA. All cells were grown in Dulbecco modified Eagle's medium (DMEM) supplemented with 10% newborn calf serum (NCS), 2mg/ml

glutamine, 100 U/ml of penicillin and 100 µg/ml streptomycin. All washes were done with PBS (Sambrook *et al.*, 1989).
Cells were maintained in a humidified incubator with 5% CO2 at 37°C.

8.3.1 RA-induction of P19 cells

To induce differentiation, P19 cells were treated with trans-retinoic acid to a final concentration of 10^{-5} M. To obtain nuclear extract the cells were collected after 6, 12, 24, 36 and 72 hours of incubation, washed with PBS, scraped and recovered by centrifugation.

8.4 Animal Experiments

8.4.1 Mouse, tissue and embryo sections

Wild type adult CD1 female mice were caged with males in the evening and monitored for the appearance of a vaginal plug the next morning. Pregnant females were sacrificed 7 to 17 d.p.c. Embryos were dissected from the maternal tissues and subsequently snap-frozen in liquid nitrogen (for RNA extraction and protein extracts) or fixed with 4% paraformaldehyde in PBS overnight at 4°C. Fixed tissues were dehydrated, paraffin embedded and sectioned at 6 µm. All animal handling conformed to the regulations issued by the ethics committee of the H.S. Raffaele.

8.4.2 DNA constructs used for mice transgenic

To analyze the *Hoxb1* enhancer, constructs, both wild-type and mutant (331 bp StuI-HindIII) r4 enhancers from the 5′ flanking region of the locus (Popperl *et al.*, 1995) were inserted into the SpeI site of the vector BGZ40, which contains the LacZ gene and SV40 poly(A) signal driven by a minimal human β-globin promoter (Yee and Rigby, 1993). Mutations in the Pbx-Meinox binding site were generated by site directed mutagenesis (TTTGTCA to cTcGTtA in association with a change of adjacent TAAT to ccgg). All mutations were confirmed by sequencing.
The minimal control *Hoxb2* enhancer construct contained a 181 bp StuI fragment; the larger *Hoxb2* construct contained a 2.1 kb BamHI-EcoRI fragment including both Krox20 dependent r3/r5 and the *Hoxb1*-dependent and r4-restricted enhancers from the flanking region of the locus inserted by blunt end ligation in the antisense orientation into the SpeI site of the BGZ40 vector (Maconochie *et al.*, 1997). 5′ and 3′ deletion variants of the 181 bp or BamHI-EcoRI enhancers were generated by PCR followed by enzymatic digestion or were generated by site-directed mutagenesis in M13 (Sculptor IVM System, Amersham). The mutation introduced into the *Hoxb2* enhancer Pbx-Meinox site converted the CTGTCA to CtcTcA and was verified by sequencing. These variants were cloned back into BGZ40 vector in a identical manner. For microinjection, inserts were separated from vector DNA by electrophoresis and purified using a gelase method.

8.4.3 Generation and analysis of transgenic mouse embryos

Transgenic embryos were generated by pronuclear injection of linearized DNA inserts into fertilized mouse eggs from an intercross of F1 hybrids (CBA X C57Bl6) followed by transfer of the injected eggs into pseudopregnant females

hosts (Whitin *et al.*, 19991). Founder embryos were harvested at the appropriate stage of development (9.5 d.p.c) and stained for LacZ reporter activity in whole mount preparation (Whitin *et al.*, 19991).

8.4.4 Generation of transgenic chick embryos

Generation of transgenic chick embryos was performed as described by Muramasu *et al* (1997) and Itasaki *et al* (1999). The regulatory regions of *Hoxb1* and *Hoxb2* were cloned into the BGZ40 vector and injected by electroporation into the neural tube of 5s-6s-7s-8s stages chick embryos. The embryos were incubated overnight before analyzing LacZ expression.

8.4.5 LacZ staining procedures

Mouse and chicken embryos (at the appropriate stage of development) were dissected from maternal tissues, washed in phosphate buffer pH 7.4 (PBS) and immediately fixed for 5-20 min, depending on developing stage, at room temperature in fix solution (2% paraformaldehyde, 0.2% glutaraldehyde, 2 mM MgCl2, 5 mM EGTA pH 7.3 in phosphate buffer 1X).
The embryos were washed 3 times for 15 min at room temperature (0.01% sodium deoxycolate, 0.02% Nonidet P40, 2 mM $MgCl_2$, in phosphate buffer 1X) and stained for one hour to overnight at 30° C or 37°C depending on the level of LacZ activity (50 mg X-gal, 0.106g potassium ferrocyanide and 0.082 g dissolved in wash buffer).

8.5 Generation of Knockout Prep1 mice

8.5.1*Prep1* targeting. Mice in which the Prep1 gene was targeted by gene trapping were obtained from a library of randomly targeted embryonic stem cell lines from Lexicon Genetics Inc. (The Woodlands, Texas, USA). This method uses random insertional mutagenesis with a fragment of DNA coding for a reporter or a selectable marker gene as mutagen. The marker genes used were beta-geo and Pur^r. Embryonic stem cells were infected with recombinant retrovirus produced from a Moloney murine leukemia virus–based packaging cell line. Puromycin-selected clones were lysed to obtain RNA to be used in the rapid amplification of cDNA ends (3′ RACE) (Zambrowich *et al.*, 1998). Direct sequencing of 3′ RACE products identified a Prep1-targeted ES cell clone (Omnibank OST71835) by comparison with sequences in the GenBank database (discussed in detail in the results). The analysis, including determination of the precise insertion site, was carried out at Lexicon Genetics.

8.5.2 Genotyping by Southern Blot. In initial studies on ES cells, determination of the genotype at the Prep1 locus was performed by using a Southern blotting method with EcoR1-restricted DNA to a Prep1-2-3 probe (see section 8.1.10).

8.5.3 PCR genotyping strategy. Using the genomic sequence, Prep1-specific oligos Prep-F1 and Prep-R1 were synthesized. Oligo Prep-R1 was used in combination with the LTR2 oligo (sequence provided below) to amplify 230 bp in the disrupted allele, while the combination of oligos Prep-F1- Prep-R1 amplify 300 bp of wild type allele.
Sequences of oligos:
Prep-F1 5′-CCAAGGGCAGTAAGAGAAGCTCTGGAG-3′
Prep-R1 5′-GGAGTGCCAACCATGTTAAGAAGAAGTCCC-3′

LTR2 5'-CAAAATGGCGTTACTTAAGCTAGCTTGCC-3'
PCR genomic screening reactions were performed using the following parameters: an initial hot-start (5 min, 95°C) followed by 5 cycles consisting of denaturation (1 min, 95°C), annealing (1 min, 58°C) and elongation (1 min, 72°C). The annealing temperature was then increased to 62°C and the amplification cycle was repeated 25 times.
A final extension step (5 min, 72°C) was performed. Genomic DNA was analyzed in a 50 µl PCR reaction. The complete sample was separated by gel electrophoresis.

8.5.4 Mice breeding. Targeted 129/SvEvBrd embryonic stem cells were injected into C57BL/6 albino blastocytes. The chimeras (129/SvEvBrd) were then crossed with C57BL/6 albinos to produce the heterozygotes (F1 cross). Mice (50% C57BL/6) heterozygous for the Prep1 disruption were mated, and their offspring were genotyped at the Prep1 locus.

CHAPTER 9
REFERENCES

Abramovich, C., Chavez, E. A., Lansdorp, P. M. and Humphries, R.. K. (2002) "Functional characterization of multiple domains involved in the subcellular localization of the hematopoietic Pbx interacting protein (HPIP)". *Oncogene* **21**:6766-71.

Abu-Shaar, M. and Mann, R. S. (1998). "Generation of multiple antagonist domains alongthe proximal axis during Drosophila leg develoment" *Development* **125**: 3821-30.

Abu-Shaar, M.Ryoo, H. D. and Mann, R. S. (1999). "Control of the nuclear localization of Extradenticle by competing nuclear import and export signal" *Genes Dev* **13**: 935-45.

Ades, S. E. and sauer, R.T. (1994). "Differential DNA-binding specificity of he engrailed homeodomain: the role of residue 50. *Biochemistry* **33**: 9187-94.

Affolter, M. marty, T. and Vigano, M. A. (1999). " Balancing import and export in development" *Genes Dev* **13**: 913-15.

Asahara, H., Dutta, S., Kao, H. Y., Evans, R. M. and Montminty, M. (1999). "Pbx-Hox heterodimers recruit co-activator-corepressor complexes in Drosophila". *Mol. Cell. Bio.* **19**: 8291-25.

Aspland, S. E. and White, R. A. (1997)."Nucleocytoplasmatic localization of extradenticle protein spatially regulated throughout development in Drosophila". *Development* **124**: 741-47.

Azpiazu, N. and Morata, G. (1998). "Functional and regulatory interactions between Hox and extradenticle genes". *Genes Dev* **12**: 261-73.

Bell, E., Wingate, R. J. and Lumsden, A. (1999) "Homeotic transformation of rhombomere identity after localized Hoxb1 misexpression". *Science.* **284**:2168-71.

Bel-Vialar, S., Itasaki, N. and Krumlauf, R. (2002) "Initiating Hox gene expression: in the early chick neural tube differential sensitivity to FGF and RA signaling subdivides the HoxB genes in two distinct groups". *Development.* **129**:5103-15

Berthelsen, J. Vandekerkhove, J. and Blasi, F. (1996). "Purifications and characterization of UEF3 a novel factor involved in the regulation of the urokinase and other Ap1 controlled promoter" *J. Biol. Chem.***271**: 3822-30.

Berthelsen, J., Kilstrup-Nielsen, C.., Blasi, F., Mavilio, F., and Zappavigna, V. (1999). "The subcellular localization of Pbx1 and Exd proteins depends on nuclear import and export signals and is modulated by association with Prep1 and Hth". *Genes Dev* **13**: 946-53.

Berthelsen, J., Zappavigna, V., Ferretti, E., Mavilio, F. and Blasi, F. (1998b). "The novel homeoprotein Prep1 modulates Pbx-Hox protein cooperativity" *EMBO J.* **17**:1434-45.

Berthelsen, J., Zappavigna, V., Mavilio, F. and Blasi, F. (1998b). "Prep1, a novel functional partner of Pbx proteins" *EMBO J.* **17**:1423-33.

Billeter M, Qian YQ, Otting G, Muller M, Gehring W, Wuthrich K. "1993" Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex. *J Mol Biol. Dec* **234**:1084-9.

Botas, J. (1993)"Control of morphogenesis and differentiation by HOM/Hox genes" *Curr Opin Cell Biol.* **5**:1015-22.

Brake, R. L., Kees, U. R. and Watt, P. M. (2002) "A complex containing PBX2 contributes to activation of the proto-oncogene HOX11". *Biochem Biophys Res Commun.* **294**:23-34.
Burling, T. and Ruvkun, G. (1992). "New motif in Pbx genes" *Nature Gen.***1**: 319-20.
Burling, T. R. (1994). A comprehensive classification of hoemobox genes "ed: D. Doboule, Oxford, UK;Oxford University press, 27-41.
Burling, T. R., (1997). "Analysis of TALE superclass homeobox genes (Meis, PBC, Knox, Iroquois, TGIF) reveals a novel domain conserved between plants and animals". *N.A.R.* **25**, 4173-4180.
Butler, A. B. and Hodos, W. (1996). "Comparative vertebrate neuroanatomy". Evolution and adaptation New York: Wiley-Liss, Inc.
Caesar, F. and Mann, R. S., (1998). " Control of antennal versus leg development in Drosophila". *Nature* **392**: 723-26.
Camus A, Tam PP. (1999). "The organizer of the gastrulating mouse embryo".
Capdevilla, J., Tsukui, T., Rodriquez Esteban, C., Zappavigna, V. and Izpisua Belmeonte, J. C. (1999). "Control of vertebrate limb outgrowth by the proximal factor Meis2 and distal antagonism of BMPs by Gremlin" *Mol Cell* **4**: 839-49.
Capovilla, M., Brandt, M., and Botas, J. (1994)" Direct regulation of decapentaplegic by ultabitorax and its role in Drosophila midgut morphogenesis". *Cell* **76**: 461-75.
Carperter, E. M., Goddard, J. M., Chisaka, O., Manley, N. R., and Capecchi, M. R. "Loss of Hoxa1 (Hox-1.6) function results in the reorganization of the murine hindbrain". *Development* **118**: 1063-75.
Carrol, S. B. (1995). Hoemotic genes and evolution of arthropods and chordates". *Nature* **376**: 479-85.
Cecconi, F., Proetetzel, G., Alvarez-bolado, G., Jay, D. and Gruss, P. (1997). "*Dev. Dynam.* **210**: 184-90.
Chan, S.-K. and Mann, R. S. (1996). "A structural model for a homeotic protein-extrandenticle DNA complex accounts for the choice of Hox proein in the heterodimer". *Proc Nat Acad Ssci USA* **93**:5223-28.
Chan, S.-K., Jaffe, L., Capovilla, M., Botas, J. and Mann, R. S. (1994). "The DNA-binding specificity of Ultrabithorax is modulated by cooperative interaction with extradenticle, an other homeoprotein". *Cell.* **78**:603-15.
Chan, S.-K., Popperl, H., Krumlauf, R., and Mann, R. S. (1996). "An extradenticle-induced conformational change in a Hox protein overcomes an inhibitory function of the conserved hexapeptide motif". *EMBO J.* **15**:2476-190.
Chan, S.-K., Ryoo, H.-D., Gould, A., Krumlauf, R., and Mann, R. (1997). "Switching the *in vivo* specificity of a minimal *HOX*-responsive element". *Development* **124**: 2007-2014.
Chang, C. P., Brocchieri, L., Shen, W. F., Largman, C., and Cleary, M. (1996). "Pbx modulation of Hox homeodoamin amino-terminal arms establishes different DNA binding specificities across the Hox locus" *Mol. Cell. Biol.* **16**: 1734-45.
Chang, C. P., de Vivo., I and Cleary, M. (1997a). The Hox cooperativity motif of the chimeric oncoprotein E2a-Pbx1 is necessary and sufficient for oncogenesis" *Mol. Cell. Biol.* **17**: 81-88.

Chang, C. P., Jacobs, Y., Nakamura, T., Jenkins, N. Copeland, N. and Cleary, M. (1997b). "mies proteins are major in vivo DNA binding partner for wild type but not chimeric Pbx proteins" *Mol. Cell. Biol.* **17**: 5679-87.
Chang, C. P., Shen, W. F., Rozenfeld, S., Lawrence, H., Largman, C., and Cleary, M. (1995). "Pbx protein display an hexapeptide-dependent cooperativite DNA bining with subset of Hox protein" *Gen. Dev.* **9**: 663-74.
Chao SH, Walker JR, Chanda SK, Gray NS, Caldwell JS. (2003) "Identification of homeodomain proteins, PBX1 and PREP1, involved in the transcription of murine leukemia virus". *Mol Cell Biol.* **23**:831-41.
Chariot, A, Gielen, J., Merville, M. P. and Bours, V. (1999). "The homeodomain-containing protein: an update on their interacting partner" *Biochem. Pharmacol.* **58**: 1851-57.
Chariot, A., van Lint, C., Chapelier, M., Gielen , J., Merville, M.P and Bours, V. (1999), "CBP and histone deacetylase inhibition enhance the transactivation potential of the HOXB7 homeodomain-containing protein". *Oncogene.* **18**:4007-14.
Chen, H. M., Rossier, C., Nakamura, Y., Lynn, A., Chakravarti, A and Antonorakis, S. E. (1997). "Cloning of a novel homeodobox-containing gene, PKNOX and mapping to human chromosome 21q22.3" *Genomics* **41**: 193-200.
Chisaka, O,. and Capecchi. M. R. (1991) "Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene hox-1.5." *Nature* **350**:473-9.
Conlon FL, Lyons KM, Takaesu N, Barth KS, Kispert A, Herrmann B, Robertson EJ. (1994)."A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse". *Development.* **120**:1919-28.
Cordes, S. P. and Barsh, G. S. (1994). "The mouse segmentation gene kr encodes a novel basic domain-leucine zipper transcription factor". *Cell* **79**:1025-34.
Couly, G., Grapin-Botton, A., Coltey, P., Ruhin, B. and Le Douarin, N.M. (1998). "Determination of the identity of the derivatives of the cephalic neural crest: incompatibility between Hox gene expression and lower jaw development". *Development* **125**:3445-59.
Couly, G., Grapin-Botton, A., Coltey, P.and Le Douarin, N. M. (1996). "The regeneration of the cephalic neural crest, a problem revisited: the regenerating cells originate from the contralateral or from the anterior and posterior neural fold". *Development* **122**:3393-407
De Cesare, D., Palazzolo, and Blasi, F. (1996). "Functional characterization of COM, a DNA region required for cooperativity between AP-1 sites in urokinase gene transcription" *Oncogne* **13**: 2551-2551.
Deshpande, A. D. Ramakrishna, W., Mulay, G. P., Gupta, V. S., and Ranjerkar, P. K. (1999). " Phylogenetic classification and molecular evolution of knotted1 homeobox genes". *Theretical& Applied Genetics* **99**: 203-09.
Di Rocco, G., Gavalas, A., Popperl, H., Krumlauf, R., Mavilio, F. and Zappavigna, V. (2001). "The recruitment of SOX/OCT complexes and the differential activity of HOXA1 and HOXB1 modulate the Hoxb1 auto-regulatory enhancer function". *J Biol Chem* **276**:20506-15.
Di Rocco, G., Mavilio, F., Zappavigna, V. (1997). "Functional dissection of a transcriptionally active, target-specific Hox-Pbx complex". *EMBO J.***16**, 3644-3654.
Doboule, D. and Morata, G. (1994). "Colinearity and functional hierarchy among genes of the homeotic complex". *Trends. Genetics* **10**:358-64

Dolle, P., Izpisua-Belmonte, J. C., Falkenstein, H., Renucci, A. and Duboule, D (1989)."Coordinate expression of the murine Hox-5 complex homoeobox-containing genes during limb pattern formation". *Nature.* **342**:767-72.
Dolle, P., Lufking, T., Krumaluf, R., Mark, M., Doboule, D., and Chambon, P. (1993). " Local alterations of Krox20 and Hox gene expression In the hindbrain suggest lack of rhombomers 4 and 5 homozygote null Hoxa-1 (Hox1.6) mutant embryos". *Proc Nat Acad Ssci USA* **90**:7666-70.
Duboule D, Dolle P. (1989) "The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes".
Duboule, D. (1995). "Vertebrate Hox genes and proliferation: an alternative pathway to homeosis?" *Curr Opin Genet Dev.* **4**:525-8.
Favier, B and Dolle', P. (1997). "Developmental function of mammalian Hox genes". *Mol. Hum.Reprod* **3**: 115-31.
Favier, B., Rijli, F. M., Frontmental-Ramain, C., Fraulob, V. Chambon, P. and Dolle', P. (1996). "Functional cooperation between the non-paralogous genes Hox10 and Hox11 in the development forelimb and axial skeleton". *Development* **122**: 449-60.
Feldman, B. J., Reid, T. R.and Cleary, M. L. (1997) "Pim1 cooperates with E2a-Pbx1 to facilitate the progression of thymic lymphomas in transgenic mice". *Oncogene.* **15**:2735-42.
Ferretti, E., Marshall, H., Popperl, E., Maconochie, M., Krumlauf, R., and Blasi, F. (2000). "Segmental expression of Hoxb2 in r4 requires two separate sites that integrate cooperative interaction between Prep1, Pbx and Hox proteins" *Development* **127**: 155-166.
Ferretti, Schulz, H., Talarico, D., Blasi, F. and Berthelsen, J. (1999). "The Pbx-regulating protein Prep1 is present in different Pbx complexed forms in mouse" *Mech. Dev.* **83**: 53-64.
Fognani, C., Kilstrup-Nielsen, C., Berthelsen, J, Ferretti, E., Zappavigna,, V. and Blasi F. (2002). "Characterization of PREP2, a paralog of PREP1, which defines a novel sub-family of the MEINOX TALE homeodomain transcription factors". *Nucleic Acids Res.* **30**:2043-51.
Frohman, M. A., dush, M. K. and martin, G. R. (1988). "Rapid amplification of full-lenght cDNA from rare transcripts: Amplification using a single specific oligonucleotide primer". *Proc Nat Acad Ssci USA* **85**:8998-9002.
Garcia-Bellido A, Ripoll P, Morata G. (1973) "Developmental compartmentalisation of the wing disk of Drosophila". *Nat New Biol.* **147**:251-3.
Gavalas A, Studer M, Lumsden A, Rijli FM, Krumlauf R, Chambon P. (1998). "Hoxa1 and Hoxb1 synergize in patterning the hindbrain, cranial nerves and second pharyngeal arch". *Development* **125**:1123-36.
Gavalas, A. (2002) "ArRAnging the hindbrain". *Trends Neurosci.* **2**:61-4.
Gavalas, A. and Krumlauf, R. (2000). "Retinoid signalling and hindbrain patterning". *Curr Opin Genet Dev* **10**:380-6.
Gavalas, A., Studer, M., Lumsden, A., Rijli, F. M., Krumlauf, R. and Chambon P. (1998). "Hoxa1 and Hoxb1 synergize in patterning the hindbrain, cranial nerves and second pharyngeal arch". *Development* **125**:1123-36.
Gavalas, A., Trainor, P., Ariza-McNaughton, L. and Krumlauf, R. (2001). "Synergy between Hoxa1 and Hoxb1: the relationship between arch patterning and the generation of cranial neural crest". *Development* **128**:3017-27.

Gehring, W. J., Affolter, M. and Burglin, T. (1994). "Homeodomain proteins" *Annu. Rev. Biochem.* **63**: 487-526.
GehringW.J., Qian, Y.Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Wurthrich, K. (1994). "Homeodomain-DNA recognition". *Cell* **78**, 211-223.
Gonsalez-Crespo, S. and Morata, G. (1996). "Genetic evidence for the subdivision of the arthropod limb into coxopodite and telopodite" *Development* **122**: 3921-28.
Gonsalez-Crespo, S., Abu-Shaar, M. torres, M., Martinez A. C., Mann, R. S. and Morata, G. (1998). "Antagonism between extradenticle function and Hedgehog signaling in the development limb" *Nature* **394**: 196-200.
Goudet, G., Delhalle, S., Biemar, F., Martial, J. A. and Peers, B. (1999). "Functional and cooperative interactions between the homeodomain Pdx1, Pbx and Prep1 factors on the somatostatin promoter" *J. Biol. Chem.* **274**: 4067-73.
Gould, A., Morrison, A., Sproat, G., White, R., and Krumlauf, R. (1997). "Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping *Hox* expression patterns". *Genes and Development* **11**, 900-913.
Graham, A. Heyman I. and Lumsden, (1993). "A. Even-numbered rhombomeres control the apoptotic elimination of neural crest cells from odd-numbered rhombomeres in the chick hindbrain". *Development* **119**:233-45.
Grammatopoulos, G. A., Bell, E., Toole, L., Lumsden, A. and Tucker, A. S. (2000). "Homeotic transformation of branchial arch identity after Hoxa2 overexpression". *Development.* **127**:5355-65.
Greer, J. M., Puetz, J. M., Thomas, K. R. and Capechi, M. R. (2000) Maintenance of functional equivalence during paralogous Hox genes evolution". *Nature* **403**: 661-665.
Gregory, P. A, and Mackenzie. P. I. (2002). "The homeodomain Pbx2-Prep1 complex modulates hepatocyte nuclear factor 1alpha-mediated activation of the UDP-glucuronosyltransferase 2B17 gene".*Mol Pharmacol.* **62**:154-61.
Guthrie S, Butcher M, Lumsden A. (1991). "Patterns of cell division and interkinetic nuclear migration in the chick embryo hindbrain". *J Neurobiol* **22**:742-54.
Haller K, Rambaldi I, Kovacs EN, Daniels E, Featherstone M. (2002) "Prep2: cloning and expression of a new prep family member". *Dev Dyn.* Nov;225(3):358-64.
Hansen, C. S., Marion, C. D., Steele, K., George, S. and Smith, W. C. (1997). "Direct neural induction and selective inhibition of mesoderm formation and epidermis inducers by Xnr3". *Development* **124**:483-92.
Herzig, S., Fuzesi, L. and Knepel W. (2000). "Heterodimeric Pbx-Prep1 homeodomain protein binding to the glucagon gene restricting transcription in a cell type-dependent manner". *J Biol Chem.* **275**:27989-99.
Hollemann T, Chen Y, Grunz H, Pieler T. (1998). "Regionalized metabolic activity establishes boundaries of retinoic acid signaling". *EMBO J.* **24**:7361-72
Huang, C. C. and Herr, W. (1996). "Differential control of transcription by homologues homeodoamin coregulatos". *Mol Cell Biol* **16**: 2967-76.
Hunt P, Gulisano M, Cook M, Sham MH, Faiella A, Wilkinson D, Boncinelli E, Krumlauf R. (1991). "A distinct Hox code for the branchial region of the vertebrate head". *Nature* **353**:861-64.

Hunt P, Krumlauf R. (1991) "Deciphering the Hox code: clues to patterning branchial regions of the head". *Cell.* **66**:1075-8.
Imoto, I., Sonoda, I., Yuki, Y.and Inazawa, J. (2001) "Identification and characterization of human PKNOX2, a novel homeobox-containing gene. *Biochem Biophys Res Commun.* **287**:270-6.
Inoue, T., and Krumlauf, R. (2001). "An impulse to the brain: Using in vivo electroporation". *Nature Neuroscience* **4** Suppl:1156-8.
Inoue, T., Nakamura, S., and Osumi, N. (2000). "Fate mapping of the mouse prosencephalic neural plate". *Dev Biol* **219,** 373-83.
Isaacs, H. V., Pownall, M. E, and Slack, J. M. (1998) "Regulation of Hox gene expression and posterior development by the Xenopus caudal homologue Xcad3". *EMBO J.* **17**:3413-27.
Itasaki, N., Bel-Viar, S., and Krumlauf, R. (1999). "Shocking developments in chick embryology; electroporation and in ovo gene expression". *Nat. Cell. Biol.*, **1** 203-207.
Izpisua-Belmonte JC, Falkenstein H, Dolle P, Renucci A, Duboule D. (1991). "Murine genes related to the Drosophila AbdB homeotic genes are sequentially expressed during development of the posterior part of the body". *EMBO J.* **10**:2279-89.
Jacobs, Y., Schnabel, C. A. and Cleary, M. (1999). "trimeric association of Hox and TALE homoeodomain proteins mediates Hoxb2 hindbrain enhancer activity" *Mol. Cell Biol.* **19**: 5134-42.
Johonson, F. B., Parker, E. and Krasnow, M. A."Extradenticle protein is a selective cofactor for the Drosophila homeotics: role of homeodomain and YPWM amino acid motif in the interaction". *Proc Nat Acad Ssci USA* **92**:739-43.
Joyner AL (1996). "Engrailed, Wnt and Pax genes regulate midbrain-hindbrain development" *Trends Genet.* **12**:15-20.
Kamps, M. P. and Baltimore, D. (1993). "E2A-Pbx1 t(1:19) translocation protein of human pre-B-cell acute lymphocytic leukemia, causes myeloid leukemia in mice". *Mol. Cell Biol.* **13**: 351-57.
Kamps, M. P. and Wright, D. D. (1994). "Oncoprotein E2A-Pbx1 t(1:19) immortalizes a myeloid progenitor in primary marrow cultures without abrogating its factor-dependence". *Oncogene* **9**: 3159-66.
Kamps, M. P. Look, A. T. and Baltimore, D. (1991). "The human t(1:19) translocation in pre B ALL produces multiple nuclear E2A-Pbx1 fusion proteins with differing transforming potentials". *Genes Dev.* **5**: 358-68.
Kamps, M. P., Murre, C., Sun, X., and Baltimore, D. (1990). "A new homeobox gene contributes the DNA binding domain of the t(1:19) translocation protein in pre-B ALL". *Cell* **60**, 547-555.
Kawagoe, H., Humphiers, R. K., Blair, A., Sutherland, H. J. and Hogge, D. E. (1999). "Expression of Hox genes, Hox cofactors, and MLL in phenotypically and functionally defined subpopulations of leukemic and normal human hematopoitic cells" *Leukemia* **13**: 687-98.
Kenyon, C. (1994). If birds can fly, why can't we? Homeotic genes and evolution". *Cell* **78**: 175-80.
Kessel M, Gruss P. (1991). "Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* **67**:89-104.

Kessel, M. and Gruss, P. (1991). "Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid". *Cell*. **67**:89-104.
Keynes, R. and Krumlauf, R. (1994). "Hox genes and regionalization of the nervous system". *Annu Rev Neurosci* **17**: 109-132.
Kilstrup-Nielsen, C., Alessio, M. and Zappavigna V. (2003) "PBX1 nuclear export is regulated independently of PBX-MEINOX interaction by PKA phosphorylation of the PBC-B domain". *EMBO J.* **22**:89-99.
Kissinger, C. R., Liu, B. S., Martin-Blanco, E., Kornberg, T. B. and Pabo, C. O. (1990). "Crystals structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understandinf homeodomain-DNA interactions" *Cell* **63**, 579-90.
Knoepfler, P. S. and Kamps, M. P. (1995). "The hexapeptide motif of Hox proteins is required for cooperative DNA binding with Pbx1, physically contacts Pbx1, and enhances DNA binding byPbx1". *Mol. Cell Biol.* **15**: 5811-09.
Knoepfler, P. S. and Kamps, M. P. (1997b). "The Pbx family of proteins is strongly upregulated by post-transcriptional mechanism during retinoid acid-induced differentiation of P19 embryonal carcinoma cell". *Mech. Dev* **63**: 5-14.
Knoepfler, P. S., Calvo, K. Y., Chen, H., Antonorakis, S. E., and Kamps, M. P. (1997a). "Meis1 and Pknox bind DNA cooperatively with Pbx1 utilizing an interaction surface disrupted in oncoprotein E2a-Pbx1". *Proc. Natl. Acad. Sci. USA* **94**, 14553-14558.
Knoepfler, P. S., Lu, Q. and Kamps, M. P. (1996). "Pbx1-Hox heterodimers bind DNA on inseparable half-sites that permit intrinsic DNA binding specificity of the Hox partner at nucleotides 3' o a TAAT motif". ". *Nuclei Acid. Res.* **24**: 2288-2294.
Kontges, G., and Lumsden, A. (1996). "Rhombenceohalic neural crest segmentation is preserved throughout craniofacial ontogeny". *Development* **122**, 3229-3242.
Krumlauf R. (1992). "Evolution of the vertebrates Hox homeobox genes". *BioEssay* **14**, 245-52.
Krumlauf R. (1994). "Hox genes in vertebrate development". *Cell* **78**, 191-201.
Krumlauf, R. (1993) "Hox genes and pattern formation in the branchial region of the vertebrate head". *Trends Genet* **9**:106-12.
Kurrant, E., Pai., C. Y. Sharf, R., Halachmi, N., Sun, Y. H. and Salzberg, A. (1998). "Dorsotonals/homotorax, the Drosophila homologue of Meis1, interacts with extradenticle in patterning of the embryonic PNS". *Development* **125:** 1037-48.
Lamb TM, Harland RM. (1995). "Fibroblast growth factor is a direct neural inducer, which combined with noggin generates anterior-posterior neural pattern". *Development.* **121**:3627-36.
Lawrence PA, Johnston P. (1989) "Pattern formation in the Drosophila embryo: allocation of cells to parasegments by even-skipped and fushi tarazu". *Development* **105**:761-7
Lawrence, H. J. and Morata, G. (1994). " Homeobox genes their function in Drosophila segmentation and pattern" *Cell* **78**:181-89.
Lawrence, H. J., Rozenfeld, S., Cruz, C., Matsukuma, K., Kwong, A., Komuves, L., Buchberg, A. M. and Largman, C. (1999). " Frequent co-expression of the

HoxA9 and Meis1 homeobox genes in human myeloid leukemias" *Leukemia* **13**:1993-09.
Le Douarin, N. M. and Ziller C. (1993). "Plasticity in neural crest cell differentiation" *Curr Opin Cell Biol.* **6**:1036-43.
LeBrun, D. and Cleary, M. L. (1994). "Fusion with E2A alters the transcriptional properties of the homeodomain protein Pbx1 in t(1:19) leukemia". *Oncogene* **9**: 1641-47.
Lemaire, P., and Kodjabachian L. (1996). "The vertebrate organizer: structure and molecules". *Trends Genet* **12**:525-31.
Lewis, E. B. (1978). "A gene complex controlling gene segmentation in Drosophila" *Nature* **276**: 565-70.
Li, X. L., Verasksa, A. and McGinnis, W. (1999). A sequence motif distinct frim Hox binding sites controls the specificity of a Hox response element" *Development* **126**: 5581-89.
Lu, Q., and Kamps, M.P. (1996a). " Selective repression of transcriptional activators by Pbx1 does not require the homeodomain" *Proc Natl Acad Sci USA* **93**, 470-80.
Lu, Q., and Kamps, M.P. (1996b). " Structural determinants within Pbx1 that mediate cooperative DNA binding with pentapeptide-containing Hox proteins: Proposal for a model of a model of a Pbx-Hox-DNA complex". *Mol. Cell. Biol.* **15**, 3786-3795.
Lu, Q., Knoepfler, P. S. Scheele, J., Wright, D. D. and Kamps, M.P. (1995). "Both Pbx1 and E2A-Pbx1 bind the DNA motif ATCAATCAA cooperatively with the product of multiple murine Hox genes, some of which are themselves oncogene". *Mol. Cell. Biol.* **15** 3786-3795.
Lu, Q., Wright, D. D. and Kamps, M.P. (1994). "Fusion with E2A converts the Pbx1 homeodomain protein into a constitutive transcriptional activator in human leukemias carrying the t(1:19) translocation". *Mol. Cell. Biol.* **14**: 3938-48.
Lufkin, T. (1999). "Transcriptional control of Hox genes in the vertebrate nervous system" *Curr Opin Genet Dev* **6**: 575-80.
Lumdsen, A. and Krumlauf, R. (1996). "Patterning the vertebrate neuraxis". *Science,* **274**: 1109-1115.
Lumsden, A., Sprawson,N.and Graham A. (1991) "Segmental origin and migration of neural crest cells in the hindbrain region of the chick embryo". *Development* **4**:1281-91
Maconochie, M., Nonchev, S., Studer, M., Chan, S. K., Popperl, H., Sham, M. H., Mann, R. S. and Krumlauf, R. (1997). "Cross-regualtion in the mouse HoxB complex: the expression of Hoxb2 in rhombomere 4 is regulated by Hoxb1" *Genes Dev* **11**:1885-95.
Manak, J., Mathies, L. D. and Scott, M. P. (1994). "regulation of decapentaplegic mid-gut enhancer by homeotic proteins" *Development,* **120**: 3605-19.
Maniatis, T., Fritsch, E. F. and sambrook, J. (1982). "Molecular Cloning: A Laboratory Manual" Cold Spring Harbor, New york: Cold Spring Harbor Laboratory Press.
Mann, R. S., (1995). "The specificity of homeotic gene function". *ByoEssays* **17**, 855-893.
Mann, R.S. and Abu-Shaar, M. (1996). " Nuclear import of the homeodomain protein extradenticle in response to WG and Dpp signaling". *Nature* **383**, 630-33.

Mann, R.S., and Affolter, M. (1998) "Hox proteins meet more partners" *Curr Opin Gen Dev.* 8: 423-429.
Mann, R.S., and Chan, S.K. (1996) "Extra specificity from extradenticle; the partnership between Hox and Pbx/Exd homeodomain proteins." *Trends. Genet.* **12**: 258-62.
Manzanares, M., Bel-Vialar, S., Ariza-McNaughton, L., Ferretti, E., Marshall, H., Maconochie, M. M., Blasi, F. and Krumlauf, R .(2001). "Independent regulation of initiation and maintenance phases of Hoxa3 expression in the vertebrate hindbrain involve auto- and cross-regulatory mechanisms". *Development* **128**:3595-607.
Manzanares, M., Cordes, S., Ariza-McNaughton, L., Sadl, V., Maruthainar, K., Barsh, G. and Krumlauf, R. (1999). "Conserved and distinct roles of kreisler in regulation of the paralogous Hoxa3 and Hoxb3 genes". *Development.* **126**:759-69.
Manzanares, M., Trainor, P. A., Nonchev, S., Ariza-McNaughton, L., Brodie, J., Gould, A., Marshall, H., Morrison, A., Kwan ,C.T., Sham, M. H., Wilkinson, D. G. and Krumlauf, R. (1999). "role of kreisler in segmentation during hindbrain development". *Dev Biol.* **211**:220-37
Mark, M., Lufkin, T., Vonesch, J., Ruberte, E., Olivo, J. C., Dolle, P., Gorry, P., Lumsden, A. and Chambon, P. (1993) "Two rhombomers are altered in Hoxa1 mutant mice." *Development.* **119**: 319-38.
Marshall, H., Morrison, A., Studer, M., Popperl, H. and Krumlauf, R. (1996). "Retinoids and Hox genes". *FASEB J.*, **10**: 969-78.
Marshall, H., Studer, M., Popperl, H., Aparicio, S., Kuroiwa, A., Brenner, S. and Krumlauf, R. (1994). "A conserved retinoic acid response element required for early expression of the homeobox gene Hoxb-1". *Nature* **370**: 567-71.
Mathis, L., Kulesa, P. M. and Fraser, S. E. (2001). "FGF receptor signalling is required to maintain neural progenitors during Hensen's node progression".*Nat Cell Biol* **3**:559-66.
McGinnis, W. and Krumlauf, R. (1992). " Homeobox genes and axial patterning" *Cell* **68**: 283-302.
McGinnis, W., Levine, M., hafen, E., Kuroiwa, A. and Gehring, W. (1984). " A conserved DNA sequence found in the homeotic genes of the Drosophila Antenna pedia and Bithorax complex" *Nature* **308**: 428-33.
McWhirter, J. R., Goulding ,M., Weiner, J. A., Chun, J. and Murre, C. (1997) "A novel fibroblast growth factor gene expressed in the developing nervous system is a downstream target of the chimeric homeodomain oncoprotein E2A-Pbx1". *Development* **124**:3221-32.
McWhirter, J. R., Neuteboom, S. T., Wancewicz, E. V., Monia, B. P., Downing, J.R.and Murre, C. (1999). "Oncogenic homeodomain transcription factor E2A-Pbx1 activates a novel WNT gene in pre-B acute lymphoblastoid leukemia". *Proc Natl Acad Sci U S A.* **96**:11464-9.
Mercander, N. Leonardo, E., Azpiazu, N., Serrano, A., Morata, G., Martinez, C., and Torres, M. (1999)." Conserved regulation of proximaldistal axis development by Meis/Hth" *Nature* **402**: 425-29.
Mikkola, I., Bruun, J-A., Holm, T. and Johansen, T. (2002). "Superactivation of Pax6-mediated transactivation from paried domain binding sites by DNA-independent recruitment of different homeodomain proteins". *J. Biol Chem.* **276**:4109-4118.

Moens, C. B. and Prince, V. E. (2002) "Constructing the hindbrain: insights from the zebrafish". *Dev Dyn.* **224**:1-17.
Moens, C. B., Yan, Y.L., Appel, B., Force, A.G. and Kimmel, C. B. (1996). "valentino: a zebrafish gene required for normal hindbrain segmentation". *Development* **122**:3981-90.
Monica, K., Galili, N., Nourse, J., Saltman, D., and Cleary, M.L. (1993). " Pbx2 and Pbx3, new homeobox genes with extensive homology to the human proto-oncogene Pbx1". *Mol Cell. Biol.* **11**: 6149-6157.
Morgan, I., In Der Rieden, I., Hooiveld, M. H. and Durston, A. J. (2000). Identifying Hox paralog groups by the Pbx-binding region" *Trends Genet* **16**: 66-67.
Morrison A. D. (1998). "1+1=r4 and much much more". *Bioessays*, **20**:794-797.
Muramatsu, T., Mizutani, Y., Ohmori, Y., and Okumura, J. " Comparison of three non viral transfection methods for foreign gene expression in early chicken embryos in ovo". (1997) *B.B.R.C.*, **230**: 376-380.
Murre, C. Mccaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Cabrera, C. V., Buskin, J. N., hauschka, S. D., Lassar, A. B., Weintrab, H. and Baltimore, D. (1989). "Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence". *Cell* **58**: 537-544.
Nakamura, T., Jenkins, N. A. and Copeland, N. G. (1996). "Identification of a new family of Pbx related homeobox genes" *Oncogen* **13**: 2235-42.
Nerlov C, De Cesare D, Pergola F, Caracciolo A, Blasi F, Johnsen M, Verde P. (1992) "A regulatory element that mediates co-operation between a PEA3-AP-1 element and an AP-1 site is required for phorbol ester induction of urokinase enhancer activity in HepG2 hepatoma cells". *EMBO J.* **11**:4573-82.
Neuteboom, S. T., and Murre, C. (1997). "Pbx raises the DNA binding specificity but not the selectivity of antennapedia Hox proteins". *Mol Cell Biol* **17**: 4696-706.
Niederreither, K. Vermot, J., Schuhbaur, B., chambon, P. and Dolle, P. (2000). Retinoic acid synthesis and hindbrain patterning in the mouse embryo" *Development* 127: 75-85.
Nieto, M. A., Bradley, L. C. and Wilkinson, D.G. (1991) "Conserved segmental expression of Krox-20 in the vertebrate hindbrain and its relationship to lineage restriction".*Development.*;Suppl **2**:59-62.
Noden D. M. (1986) "Patterning of avian craniofacial muscles". *Dev Biol.* **6**:347-56
Nonchev, S., Vesque, C., Maconochie, M., Seitanidou, T., Ariza-McNaughton, L., Frain, M., Marshall, H., Sham, M. H., Krumlauf, R. and Charnay, P. (1996). "Segmental expression of Hoxa-2 in the hindbrain is directly regulated by Krox-20". *Development* **122**:543-54.
Nourse, J., Mellentin, J. D., Galili, N., Wilkinson, J., Stanbridge, E., Smith, S. D., Cleary, M. (1990). " Chromosomal translocation t(1:19) results in synthesis of a homeobox gene involved in myeloid leukemia in BXH-2 mice". *Mol. Cell. Biol.* **15**: 5434-5443.
Osumi, N., and Inoue, T. (2001). "Gene transfer into cultured mammalian embryos by electroporation". *Methods* **24**: 35-"42.
Oulad-Ablelghani, M., Chazaud, C., Bouillet, P., Sapin, V., Chambon,, P. and Dolle, P. (1997). " Meis2, a novel mouse Pbx-related homeobox gene induced by retinoc acid during differentiation of P19 embryonal carcinoma cells". *Dev Dyn.* **210**: 173-83.

Pai, C.. Y., Ku, T. S., jaw, T. J., Kurant, E., Chen, C. T., Bessarab, D. A. Salzberg, A. and Sun, Y. H. (1998). " The homothorax homeoprotein activates the nuclear localization of an other hoemoprotein, extradenticle and suppresses eye development in Drosophila" *Genes Dev.* **12:** 435-46.
Palazzolo, M., Berthelsen, J., De Cesare, D. and Blasi F. (2000). "Oct-1 specifically binds the UEF4 site of the human AP1-regulated urokinase enhancer". *Eur J Biochem.* **267**:5427-37
Paro, R. (1990) "Imprinting a determined state into the chromatin of Drosophila". *Trends Genet.* **12**:416-21.
Paro, R. (1993) "Mechanisms of heritable gene repression during development of Drosophila". *Curr Opin Cell Biol.* 6:999-1005.
Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S. and Aggarwal, A. K. (1999). Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex" *Nature* **397**: 714-9.
Peltenburg, L.T. and Murre, C. (1996). "Engrailed and Hox homeodomain proteins contain a related Pbx interaction motif that recognizes by a conserved peptide motif". *EMBO J.* **15**: 3385-93.
Peltenburg, L.T. and Murre, C. (1997). "Specific residues in the Pbx homeodomain differentially modulate the DNA-binding activity of Hox and Engrailed proteins". *Development* **124**: 1089-98.
Penkov, D., Tanaka, S., Di Rocco , G., Berthelsen, J., Blasi, F.and Ramirez, F. (2000). "Cooperative interactions between PBX, PREP, and HOX proteins modulate the activity of the alpha 2(V) collagen (COL5A2 promoter". *J Biol Chem* **275**:16681-9.
Pfeifer, M. and Wieschaus, E. (1990). "Mutations in Drosophila gene extradebnticle affect the way specific homeodomain proteins regulate segmental identity" *Genes Dev* **4**: 1209-23.
Phelan, M.L. and Featherstone, M.S. (1997) "distinct Hox N-terminal arm residues are responsible for specificity of DNA recognition by Hox monomer and Hox-Pbx heterodimers". *J Biol Chem.***272**: 8635-43.
Phelan, M.L., Rambaldi, I., and Featherstone, M.S. (1995) "Cooperative interactions between Hox and Pbx proteins mediated by a conserved peptide motif". *Mol. Cell. Biol.***15**: 3989-3997.
Pinsonneault, J., Florence, B., Vaessin ,H. and McGinnis W. (1997) "A model for extradenticle function as a switch that changes HOX proteins from repressors to activators". *EMBO J.* **16**:2032-42.
Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L. and Wolberger, C. (19990" Structure of a Hoxb1-Pbx1 heterodimer bound to DNA:role of the hexapeptide and a fourth homeodoman helix in complex formation. *Cell*, **96**: 587-97.
Pirrotta, V. (1997). *PcG* complexes and chromatin silencing. *Curr. Opin. Genet. Dev.* **7**: 249-258.
Popperl, H., Bienz, M., Studer, M., Chan, S.K., Aparicio, S., Brenner, S., Mann, R. S. Krumlauf, R. (1995). "Segmental expression of Hoxb1 is controlled by a highly conserved autoregolatory loop dependent upon exd/pbx". *Cell*, **81**: 1031-1042.
Popperl, H., Rikhof ,H, Chang,H., Haffter, P., Kimmel, C.B, and Moens, C. B. (2000)."lazarus is a novel pbx gene that globally mediates hox gene function in zebrafish". *Mol Cell.* **6**:255-67.

Prince, V. and Lumsden, A. (1994). "Hoxa-2 expression in normal and transposed rhombomeres: independent regulation in the neural tube and neural crest". *Development*.**120**: 911-23.
Qian, Y. Q., Otting, G., Billeter, M., Muller, M., Gehring, W. and Wuthrich, K. (1993) "Nuclear magnetic resonance spectroscopy of a DNA complex with the uniformly 13C-labeled Antennapedia homeodomain and structure determination of the DNA-bound homeodomain". *J Mol Biol.* **87**:1070-83.
Qian, Y., Billeter, M., Otting, G., Muller, M., Gehing, W. J. and Wuthrich, K. (1989). "The structure of the Antennapedia homoromain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors" *Cell*, **59**: 573-80.
Rauskolb, C. and Wieschaus, E., (1994). "Coordinate regulation of downstream genes by extradenticle and the homeotic selector proteins". *EMBO J*, **13**: 3561-69.
Rauskolb, C. Peifer, M., Wieschaus, E., (1993). "Extradenticle, a regulator of homeotic gene activity, is a homolog of the homeobox containing human protoncogene pbx1". *Cell*, **74**; 1017-1112.
Rauskolb, C., Smith, K. M., Peifer, M. and Wieschaus, E., (1995). "extradenticle determines segmental identities throughout Drosophila development". *Development* **121**: 3663-73.
Redmond, L., Hochfield, S. and Morabito, M. A. "The divergent homeobox gene Pbx1 is expressed in the postnatal subventricular zone and inerneurons on the olfactory bulb" *J Neurosci* **16**: 2972-82.
Rieckhof, G.E., Casares, F., Ryoo, H.D., Abu-shaar, M., and Mann, R.S. (1997). "Nuclear translocation of Extradenticle requires homothorax which encodes an extra-denticle related homeodomain protein". *Cell* **91**:171-183.
Rijli, F., Gavalas, A. and Chambon, P. (1998). Segmentation and specification in the branchial region of the head: The role of *Hox* selector genes. *Int. J. Dev. Biol.* **42**: 393-401.
Rossel, M.and Capecchi, M. R. (1999) "Mice mutant for both Hoxa1 and Hoxb1 show extensive remodeling of the hindbrain and defects in craniofacial development". *Development* **126**:5027-40.
Ryan, A. K. and Rosenfeld, M. G. (1997). "POU domain family values: flexibility, partnerships, and developmental codes". *Genes Dev.* **11**:1207-25.
Ryoo, H. D. and Mann, R. S. (1999). The control of trunk Hox specificity and activity by Extradenticle. *Genes Dev* **13**: 1704-16.
Ryoo, H. D., Marty, T., Casares, F., Affolter, M., and Mann, R. S. (1999). "Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex". *Development* **126**: 5137-5148.
Saldivar, J., Krull, C., Krumlauf, R., Ariza-McNaughton, L. and Bronner-Fraser, M. (1996). Rhombomere of origin determines autonomous versus environmentally regulated expression of *Hoxa3* in the avian embryo. *Development* **122**: 895-904.
Saleh, M., Rambaldi, I., Yang, X. J. and Featherstone MS. (2000). "Cell signaling switches HOX-PBX complexes from repressors to activators of transcription mediated by histone deacetylases and histone acetyltransferases". *Mol Cell Biol.* **20**:8623-33.

Sambrook, J., Fritsch, E., and Maniatis, T. (1989). "Molecular cloning: A Laboratory manual" Cold Spring Harbor, New York: Cold Spring harbpr laboratory Press.
Sanger, F., Nicklen, S., and Coulson, A. R. (1997). DNA sequencing with chain terminating inhibitors" *Proc Natl Acad Sci USA* **74**: 5463-67.
Schnabel CA, Selleri L, Jacobs Y, Warnke R, Cleary ML. (2001). "Expression of Pbx1b during mammalian organogenesis". *Mech Dev.* 100:131-5
Scholer HR. (1991). " Octamania: the POU factors in murine development".*Trends Genet* 7:323-9
Schumacher, A. and Magnuson, T. (1997). Murine *Polycomb*- and *trithorax*-group genes regulate homeotic pathways and beyond. *TIG* **13**: 167-170.
Scott, M P. (1999). "Hox proteins reachout round DNA" *Nature* **397**, 649-51.
Scott, M. P. and Weiner, A. J. (1984)."Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila". *Proc Natl Acad Sci U S A.* **81**:4115-9.
Selleri, L., Depew, M.J., Jacobs, Y., Chanda, S.K., Tsang, K. Y, Cheah, K.S.E. Rubenstein, J.L.R., O'Gorman, S. and Cleary, (2001). "Requirement for Pbx1 in skeltal patterning and programming chindrocyte proliferation and differentiation". *Development* **128**: 3543-3557.
Sham, M. H., Vesque, C., Nonchev, S., Marshall, H., Frain, M., Das Gupta, R., Whiting, J., Wilkinson, D., Charnay, P. and Krumlauf, R. (1993). The zinc finger gene *Krox-20* regulates *Hoxb-2 (Hox2.8)* during hindbrain segmentation. *Cell* **72**: 183-196.
Sham, M.-H., Hunt, P., Nonchev, S., Papalopulu, N., Graham, A., Boncinelli, E. and Krumlauf, R. (1992). Analysis of the murine *Hox-2.7* gene: conserved alternative transcripts with differential distributions in the nervous system and the potential for shared regulatory regions. *EMBO J.* **11**: 1825-1836.
Shanmungam, K., Featherstone, M. S., Saragovi, H. U. (1999). "Residues flanking the Hox YPWM motif contribute to cooperative interactions with Pbx". *J. Biol Chem.***272**: 19081-87.
Shanmungam, K., Green, N. C., Rambaldi, I., Saragovi, H. U. and Featherstone, M. S. (1999). " Pbx and Meis as non-DNA binding partners in trimeric complex with Hox proteins". *Mol. Cell. Biol.***19**: 7577-88.
Sharkey, M., Graba, Y. and Scott, M. P. (1997)."Hox genes in evolution: protein surfaces and paralog group" *Trends. Genetics.* **13**: 145-151.
Sharpe, J., Nonchev, S., Gould, A., Whiting ,J. and Krumlau,f R. (1998). "Selectivity, sharing and competitive interactions in the regulation of Hoxb genes". *EMBO J* **17**:1788-98.
Shen, W. F., Krishnan, K., Lawrence, H. J.and Largman, C. (2001). "The HOX homeodomain proteins block CBP histone acetyltransferase activity". *Mol Cell Biol* **21**:7509-22.
Shen, W. F., Montgomery, J., Rozenfeld, S., Moskow, J. J., Lawrence, H. J. Buchberger, A. M. and Largaman, C. (1997a). AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins". *Mol. Cell. Biol.***17**: 6448-58.
Shen, W. F., Rozenfeld, S., Lawrence, H. J. and Largaman, C. (1997b). The AbdB-like homeodomain proteins can be subdivided by the ability to form complexes with Pbx1a on a novel DNA target". *J. Biol Chem.***272**: 8198-206.

Simeone, A., Acampora, D., Arcioni, L., Andrews, P. W., Boncinelli ,E.and Mavilio, F. (1990). "Sequential activation of HOX2 homeobox genes by retinoic acid in human embryonal carcinoma cells". *Nature.* **346**:763-6.
Smith, J. E., Afonjia, O., Yee, H.T., Inghirami, G. and Takeshita, K. (1997a). "Chromosomal mapping to 15q14 and expression analysis of the human Meis2 homeobox gene" *Mamm Genome* .**8**: 951-52.
Smith, J. E., Bollekens, J. A., Inghirami, G. and Takeshita, K. (1997b). "Cloning and mapping of the Meis1 gene, the human homolog of a murine leukemogenic gene" *Genomics* .**43**: 99-103.
Smoth, W. C., and Slack, J. M. W. (1992). "dorsalization and neural induction: properties of the organizer in Xenopus Laevis. *J. Embryol. Exp. Morphol.* **78**: 299-317.
Sprules T, Green N, Featherstone M, Gehring K., (2000). "Conformational changes in the PBX homeodomain and C-terminal extension upon binding DNA and HOX-derived YPWM peptides"*Biochemistry*.**39**:9943-50.
Sprules, T., Green, N., Featherstone, M. and Gehring, K. (2003). "Lock and key binding of the HOX YPWM peptide to the PBX homeodomain". *J Biol Chem.* **278**:1053-8.
Steelman, S., Moskow, J. J., Muzynski, K., North, C., Druck, T., Montgomery, J. C., Huebner, K., Daar, I. O. and Buchberger, A. M. (1997)."Identification of a novel family of Meis1-related homeobox gene". *Genom Res* .**7**: 142-56.
Sthrul, G. (1981). "A homeotic mutation transforming leg to antenna in Drosophila". *Natures* .**292**: 635-8.
Streit, A., Berliner, A. J., Papanayotou, C., Sirulnik, A.and Stern C. D..(2000). "Initiation of neural induction by FGF signalling before gastrulation". *Nature.* **406**:74-8
Struhl, G. and Akam M. (1985). "Altered distributions of Ultrabithorax transcripts in extra sex combs mutant embryos of Drosophila". *EMBO J.*;**4**:3259-64.
Studer, M., Gavalas, A., Marshall, H., Ariza-McNaughton, L., Rijli, F., Chambon, P., and Krumlauf, R. (1998). " Genetic interaction between Hoxa1 and Hoxb1 reveal new roles in regulation of early hindbrain patterning". *Development*, **125**: 1025-1036.
Studer, M., Lumsden, M., Ariza-McNaughton, L., Bradley, A., and Krumlauf, R. (1996). " Altered segmental identity and abnormal migration of motor neurons in mice lacking Hoxb1. *Nature*, **384**: 630-634.
Studer, M., Popperl, H., Marshall, H., Kuroiwa, A. and Kumlauf, R. (1994) "Role of a conserved retinoic acid response element in rhombomere restriction of Hoxb-1". *Nature* **370**:567-71
Sun, B., Hursh, D. A., Jackson, D. and Beachy, P. A. (1995). " Ultrabithorax protein is necessary but not sufficient for full activation of decapentaplegic expression in the visceral mesoderm". *EMBO J.***14**: 520-35.
Swift, G. H., Liu, Y., Rose, S. D., Bischof, L. J.,Steelman, S., Buchberger, A. M., Wright, C. V. and Mac Donald, R. J. (1998). "An endocrine-exocrine switch in the activity of the pancreatic homeodomain protein Pdx1 through formation of a trimeric complex Pbx1b and Mrg1 (Meis2)". *Mol. Cell. Biol.***18**: 5109-20.
Swindell, E. C., Thaller, C, Sockanathan, S., Petkovich, M., Jessel, T.M. and Eichele, G. (1999). " Complementary domaims of retinoic acid production and degradation in the early chick embryo". *Dev. Biol.* **216**, 282-296

Takihara Y, Tomotsune D, Shirai M, Katoh-Fukui Y, Nishii K, Motaleb MA, Nomura M, Tsuchiya R, Fujita Y, Shibata Y, Higashinakagawa T, Shimada K. (1997). 'Targeted disruption of the mouse homologue of the Drosophila polyhomeotic gene leads to altered anteroposterior patterning and neural crest defects".*Development* **124**:3673-82.
Tanabe, Y., and Jessell T. M. (1996)."Diversity and pattern in the developing spinal cord". *Science.* **15**: 1115-20.
Thorsteinsdottir, U., Krosl, J., Kroon, E., Haman, A., Hang, T. and Sauvageau, G. (1999). "The oncoprotein E2A-Pbx1a collaborated with Hoxa9 to acutely transform primary bone marrow cells" *Mol. Cell. Biol.***19**: 6355-66.
Tjian R, and Maniatis T.(1994). "Transcriptional activation: a complex puzzle with few easy pieces".*Cell.* 77:5-8.
Toresson, H., Parmar, M.and Campbell, K. (2000) "Expression of Meis and Pbx genes and their protein products in the developing telencephalon: implications for regional differentiation". *Mech Dev.* **94**:183-7.
Trainor, P. A. and Krumlauf, R. (2000). "Patterning the cranial neural crest: hindbrain segmentation and Hox gene plasticity". *Nat Rev Neurosci.* **1**:116-24.
Trainor, P. A. and Krumlauf, R. (2001). "Hox genes, neural crest cells and branchial arch patterning". *Curr Opin Cell Biol.* **13**:698-705.
Trainor, P. A., Ariza-McNaughton, L. and Krumlauf, R. (2002). "Role of the isthmus and FGFs in resolving the paradox of neural crest plasticity and prepatterning". *Science.* **295**:1288-91.
Treismann, J., Gonczy, P., Vashishtha, M., Harris, E. and Desplan, C. (1989). "A single amino acid can determine the DNA binding specificity of homeodoman proteins". *Cell.***59**: 553-62.
Urbanek, P., Wang, Z. Q., Fetka, I., Wagner, E. F, Busslinger, M. (1994). "Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP". *Cell.* **79**:901-12.
van der Vliet PC, Verrijzer CP. (1993). "Bending of DNA by transcription factors". *Bioessays.* **15**:25-32.
Van Dijk, M. A., and Murre, C. (1994). "extradenticle raises the DNA binding specificity of homeotic selector gene products". *Cell.* **78**: 617-24.
Van Dijk, M. A.,Peltenburg, L. T. and Murre, C. (1995). "Hox gene products modulate the DNA binding activity of Pbx1 and Pbx2". *Mech. Devel.* **52**: 99-108.
Van Dijk, M. A.,Voorhoeve, P. M. and Murre, C. (1995). "Pbx1 is conserved into a transcriptional activator upon acquiring the N-teminal region of E2Ain pre-B-cell acute lymphoblastoid leukemia". *Proc Natl Acad Sci USA* **90**: 6061-65.
Vesque, C., Maconochie, M., Nonchev, S., Ariza-McNaughton, L., Kuroiwa, A., Charnay, P. and Krumlauf, R. (1996). *Hoxb-2* transcriptional activation by *Krox-20* in vertebrate hindbrain requires an evolutionary conserved *cis*-acting element in addition to the Krox-20 site. *EMBO J.* **15**:5383-5896.
Wagner, K., Mincheva, A., Korn, B., Lichter, P.and Popperl H. (2001). "Pbx4, a new Pbx family member on mouse chromosome 8, is expressed during spermatogenesis". *Mech Dev.* **103**:127-31.
Waskiewicz, A. J., Rikhof, H. A. and Moens, C. B. (2002). "Eliminating zebrafish pbx proteins reveals a hindbrain ground state". *Dev Cell.* 3:723-33.
Wassef M, Joyner AL. (1997) "Early mesencephalon/metencephalon patterning and development of the cerebellum". *Perspect Dev Neurobiol* **5**: 3-16.
White, R. (1994). " homeotic genes seek partner". *Curr Opin Biothech.* **7**:48-50

Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P. W. J., Stott, D. and Allemann, R. K. (1991). Multiple spatially specific enhancers are required to reconstruct the pattern of *Hox-2.6* gene expression. *Genes Dev.* **5**:2048-2059.
Wilkinson, D. G., Bhatt, S., Chavrier, P., Bravo, R. and Charnay, P. (1989a). Segment-specific expression of a zinc-finger gene in the developing nervous system of the mouse. *Nature* **337**:461-465.
Wilkinson, D. G., Bhatt, S., Cook, M., Boncinelli, E. and Krumlauf, R. (1989b). Segmental expression of Hox-2 homeobox-containing genes in the developing mouse hindbrain. *Nature* **341**: 405-409.
Wilson, D. and Desplan, C. (1995). "Homeodomain proteins. Cooperating to be different *Curr Biol* **5**: 32-34.
Wilson, D. and Desplan, C. (1999). "Structural basis of Hox specificity" *Nature Structural Biology* **6**: 297-300.
Wilson, D., Sheng, G., Lcuit, T., Dostatni, N. and Desplan, C. (1993). "cooperative dimerization of paried class homeodomains on DNA". *Genes Dev* **7**: 2120-34.
Wizenmann, A. and Lumsden, A. (1997). "Segregation of Rhombomeres by Differential Chemoaffinity". *Mol Cell Neurosci* **9**:448-59.
Wolberger, C., Pabo, C. O., Vershon, A. K. and Johnson AD. (1991). "Crystallization and preliminary X-ray diffraction studies of a MAT alpha 2-DNA complex". *J Mol Biol.* **217**:11-3.
Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. and Pabo, C. O. (1991). "crystal structure of a MAT alpha2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions *Cell* **67**: 517-28.
Wolberger, C., Vershon, A. K., Liu, B., Johnson, A.D. and Pabo, C. O. (1991). "Crystal structure of a MAT alpha 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions". *Cell.* **67**:517-28.
Wu, J., and Cohen, S. M. (1992). "Proximal distal axis formation in the Drosophila leg: subdivision into proximal and distal domains by homothorax and distal less". *Development* **126**:109-17.
Yokouchi, Y., Sasaki, H. and Kuroiwa, A. (1991). "Homeobox gene expression correlated with the bifurcation process of limb cartilage development". *Nature* **353**:443-5.
Zeng, C., Pinsonneault, J., Gellon, G., McGinnis, N. and McGinnis, W. (1994). Deformed protein binding sites and cofactor binding sites are required for the function of a small segment-specific regulatory element in Drosophila embryos. *EMBO J.* **13**: 2362-2377.
Zhang, M., Kim, H.-J., Marshall, H., Gendron-Maguire, M., Lucas, A. D., Baron, A., Gudas, L. J., Gridley, T., Krumlauf, R. and Grippo, J. F. (1994). Ectopic *Hoxa-1* induces rhombomere transformation in mouse hindbrain. *Development* **120**: 2431-2442.
Zo, L.I., Dorfman, D. M. and Orkin, S. H. (1989). "The polymerase chain reaction colony miniprep". *BioTechnique* **7**:61-65.

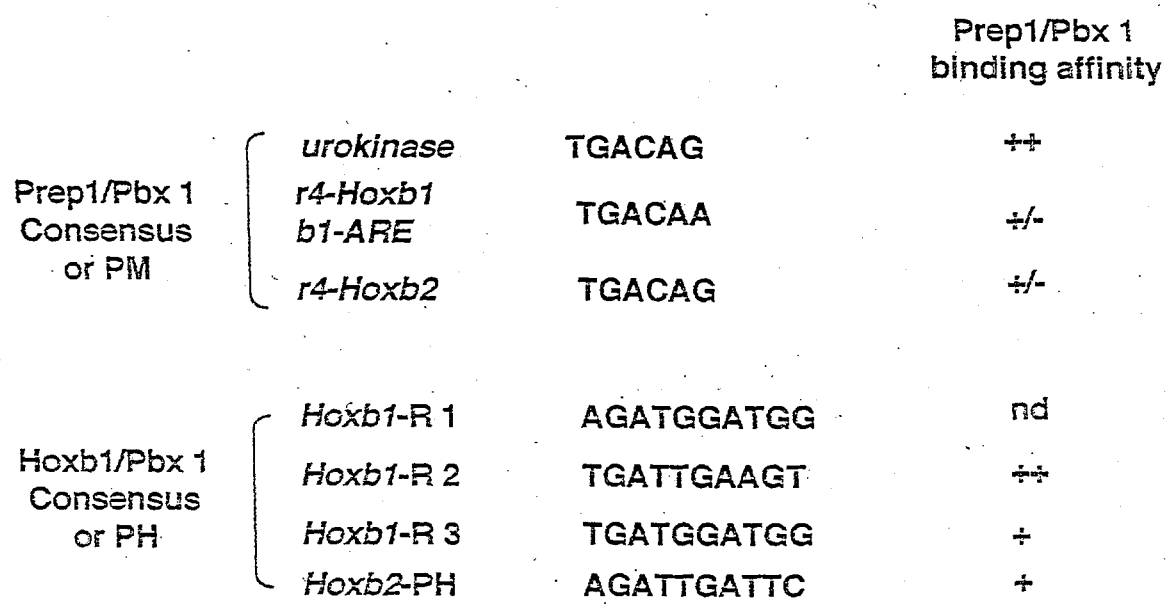


Figure 7.1
The Pbx-Meinox binding sites (PM) of *urokinase*, *Hoxb1* and *Hoxb2* enhancers show a different binding specificity for Prep1-Pbx complexes. Only the PM site in the *urokinase* enhancer bind Prep1-Pbx dimers with high affinity.
In addition, the Pbx-Hox binding sites (PH) of *Hoxb1* and *Hoxb2* enhancers also bind the Prep1-Pbx dimer.
The comparison of PM and PH binding sites suggest that the new consensus for the binding of Prep1-Pbx dimers is T/AGANNG.